UNIVERSITY OF TARTU

INSTITUTE OF GERMANIC, ROMANCE AND SLAVONIC

LANGUAGES AND LITERATURES

EMPLOYING TECHNICAL MEANS FOR ENSURING CONSISTENT TRANSLATION OF RECURRING TERMS

MA thesis

PIRET JÕGIS

SUPERVISOR: ASST. LECT. PIRET RÄÄBUS

TARTU

2013

ABSTRACT

The motivation for writing this paper had two igniters. First was the realisation of the fact that in very many kinds of translations a situation where a recurring source text word or a string of words has received more than one equivalent in the target text can be considered a grave mistake, one that may also have serious consequences. But also that such mistakes slip in however highly valued or generally knowledgeable the translator is. This inspired the author to find solutions and ways how translators can more easily avoid making such mistakes in the future.

The other igniter was a wish to understand better how computer assisted translation tools are and can be used for making a translator's work easier. Which software do translators prefer and why? How can these be best employed? For the purpose of this paper software tools were only viewed through the spectrum of the problem with inconsistent translation of recurring terms. Therefore, the thesis analyses how translators' choices in using tools helps them to avoid giving terms different equivalents.

The aim of the thesis was to find out how different software can aid the translator in the task of ensuring consistent translation of recurring terms.

The first chapter of the paper introduces the subject matter, which is illustrated by relevant examples. Also, it discusses the types of computer assisted translation tools associating these with term translation issues.

In order to gain more than just anecdotal evidence on translator behaviour, a study was carried out amongst working translators who were contacted through translation bureaus registered in Estonia. Aspiring translators studying at the University of Tartu and at Tallinn University were also added to the group of people who were asked to fill in the study questionnaire. The aim of the study was to receive more information about the translators' computer use and background, as well as their behaviour in translating texts where the inconsistent translation of terms would be considered a translation mistake. For that purpose the respondents were presented with a particular translation case and asked about their work choices within that frame. Also, in order to extend the study and to be able to draw more specific conclusions, the respondents were asked about their interest and need concerning computer studies. The second chapter of the thesis presents and discusses the study and its results, and draws conclusions thereof.

In the third chapter the author demonstrates the possibilities of computer assisted translation tools with selected programs that are known to be used by translators. The aim of these demonstrations is to see how a translator can pay attention to term translation consistency working with these programs without having to change their work methods too much. Thereby it was possible to offer translators ideas on how to structure their work differently or what to change in their work practices to be more knowledgeable in ensuring term consistency.

The paper consists of 110 pages, includes a total of 34 charts and figures, and it has three appendices.

CONTENTS

LIST OF ABBREVIATIONS

CAT – computer-assisted translation or computer-aided translation

SL – source language

ST – source text, the original text to be translated

TL – target language

TT – target text, the translated text

INTRODUCTION

All professional translators in the 21st century use electronic tools and resources in their work to some extent. In an age of ever-accelerating technological advancements where the interrelationship between translators and technology is only deepening (O'Hagan 2013: 503) it is important to learn how translators can best employ this relationship and take advantage of the technical tools they use. Some may suggest that the best way forward is to use better and smarter software and, although, in some cases that may hold true, translators can still only operate using the tools that are available for them at a given time. Therefore, having the skills and knowledge that allows translators to make the best use of the tools they already use is essential. Computer literacy, not only ability to use software for our desired purpose, but also knowledge of how to make electronic tools and resources most efficient in cooperation with the user is becoming increasingly more important and that knowledge should be based on awareness of available options.

Within this thesis a survey was carried out amongst professional and trainee translators in which they were asked, among other things, to elaborate on their software use habits. Based on that survey, a limited number of programs were selected for demonstrations that were to enable the analysis of how these programs facilitate achieving better results. In order to limit the analysis, a particular aspect was concentrated on. That chosen aspect is the consistent translation of recurring terms, which has particular importance in non-fictional texts, e.g. legal, scientific or technical texts, or, in order to avoid the pitfall of categorizing text in accordance with situational criteria: informative, argumentative or instructional texts.

Therefore, the aim of this thesis is to find out how different software can aid the translator in the task of ensuring consistent translation of recurring terms. The objective is to find ways

how to transfer more mechanical translation tasks – such as memorising which words in the text are to be treated as terms, and which are their chosen equivalents – to the computer, so that the translator would get maximum help in making certain that recurring terms receive only one equivalent and that the translator could concentrate more on other aspects of translation.

The first chapter of this paper elaborates on the need of consistent translation of terms and gives relevant examples of respective translation mistakes. The chapter continues to discuss how recurring terms are or should be treated. In which stage of the translator's work do terms get their equivalents and how are they applied? To which extent does a translator have to depend on their own personal memory? The first chapter also includes an introduction of commonly used translation tools and how these operate to facilitate the consistent translation of recurring terms.

That leads to the survey that was carried out in Estonia in February 2013. The aim of the survey was to find out translators' habits concerning software use and term translation – whether terms are treated before starting to translate or during translating, which programs are used and why. The second chapter gives detailed information about the survey, its results and conclusions.

The third chapter aims to analyse selected programs as examples of their kind and see how these programs can be used to achieve better consistency in term translation. The choice of programs was based on the author's own experience and on the study results, and included a word processor, a translation memory tool, and a terminology management tool. Also, a corpus analysis tool is demonstrated to see whether this could facilitate the translator's work in ensuring term consistency. The analysis of the programs is based on a test translation that was

chosen by the author keeping in mind that the document should not only be the kind that would insist on consistent translation of terms within itself, but also in the broader text environment. The demonstrations are accompanied by many screenshots to give a better idea of the translation environments to those not familiar with the programs.

Hereby, it is important to mention that the author does not claim to be a professional user of any of the covered programs and has not passed extensive study programmes covering all of the discussed tools. She has, however, intermediate-level daily relevant software user skills developed through various short-term computer courses, both in-class and online, and years of experience in using the same. Therefore, despite the fact that the author has made an effort to provide the readers with practical and helpful information, there may be some commands or functions she is unaware of and has therefore not considered. Many displayed methods have been learnt through personal contacts with colleagues or from professors, some were revealed to the author by people participating in the study, for which the author would hereby like to thank all of the named. Also, thanks go to the supervisor of the thesis, Piret Rääbus, for the encouragement and relevant, helpful comments and to Kaspars Kļaviņš from SIA Mestako for cooperation. In connection to the latter, the author openly admits having been in contact with and helped by a certain software provider. However, it is pertinent to stress that the aim of this thesis is not to advertise any software in question. All analysis and opinions in the paper are honest and the author has not been offered any kind of personal or other benefit in order to make any tools, programs or software look better in any way.

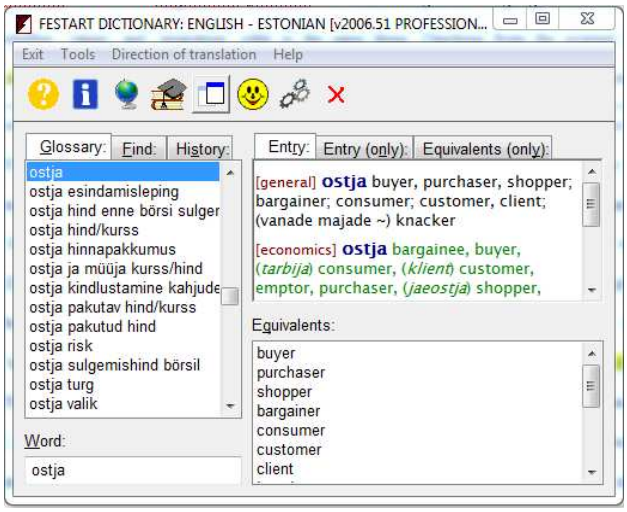
Before going further, there are a few terms applied throughout this paper with a particular, in some cases restricted meaning. Their definitions for the purposes of this paper are as follows:

- Term – any word or string of words that has a particular preferred translation equivalent in the text environment and should be translated in a consistent manner in order to avoid making a translation mistake;
- Termbase – a glossary; a database that includes selected words from the ST and their preferred equivalents in the TL;
- Translation environment – the full translated text, plus all relevant source documents that are available in the languages in question, particularly in the TL. These sources may include but are not limited to: governing law, client's web page, documents referred to in the translated document;
- Translation mistake – an error similar to the one described in this paper where inconsistent use of a term is present in the TT but not in the ST.

1. CONSISTENT TRANSLATION OF TERMS AND WHY IT MATTERS

In literature we welcome, even desire, variety and synonyms, but when it comes to other kinds of texts, e.g. legal, scientific or technical texts, we are not as tolerant. At best the use of synonyms, or inconsistent translation of the same word, may go unnoticed, but it may also bring about confusion or even disputes. Katre Kasemets stated in a lecture held at Tartu University in 2012, talking of legal text requirements that, “Synonyms are to be avoided at all costs – terms must be translated uniformly independent of the fact that they might have 100% synonyms.” The following example may help to explain why it is hard to argue against such a strong statement.

The Estonian word ‘ostja’ can be translated into English as ‘buyer’, ‘purchaser’, ‘customer’, or ‘client’, to name a few (see Example 1), all of which can be considered acceptable, depending on the context.



Example 1: English language equivalents to the Estonian word ‘ostja’ (Festart 2006).

There could be a contract where ‘ostja’ has been defined as ‘client’, but in the further text there is a clause whereby some responsibility has been laid to a ‘buyer’. Should the translated

version be signed as the original document, the ‘client’ may later claim not having the responsibility assigned to the ‘buyer’, since they have been defined as the ‘client’. If no other legal remedies are available, this situation may, in extreme cases, lead to an expensive legal dispute for which the translator really has responsibility (provided the original document indeed used ‘ostja’ throughout). As professional translators we do not only want to avoid similar situations, but want to provide our clients with translations that would not include any such mistakes, whether these lead to misunderstandings or not.

As was displayed by the example, it must be made sure that any term that appears in the ST even just twice, is translated into the TT in the same way in each occasion. Yet, considering the translation environment, one must also pay attention to terms that may only appear in the ST just once. Otherwise we might find ourselves in a situation where a contract governed by the Law of Obligations Act of the Republic of Estonia includes the word ‘enampakkumine’ and the translator of the contract has used ‘bidding’ or ‘tender’ as its equivalent, whereas the translated version of the Law of Obligations Act available in the Estonian state gazette, Riigi Teataja, uses the equivalent ‘auction’ (Legaltext 2013). Should a reader be in a situation where they need to compare the contract conditions to the provisions of the named act, it might be difficult to ascertain whether ‘bidding’ or ‘tender’ is the exact equivalent of ‘auction’ or not.

However, there is more reason to follow Kasemets’ words. There are many environments in which the translation equivalents of words are not free for the translator to choose – this may include organisational or thematic restrictions – there might be relevant standardisation or organisational guidelines in place. Failing to use such an equivalent and opt for a synonym that may content-wise be correct, may confuse readers, lead to misunderstandings, etc., or even anger the client. Arvi Tavast (2008: 106–111) has studied what clients ordering

translation evaluate about a translation and found that half of the studied clients named inconsistency in translated terms as one of the main error types; where 'inconsistency' means that the translator has "ignored dictionaries provided by the customer or the terminology usage established in the target language, or the use of terms is inconsistent within the text." So searching for the equivalents of terms paying attention to the translation environment and staying client-specific – following specialist dictionaries, style guides, or standards that may be provided by the customer – may enable to avoid another situation of confusion and distress. There are many arguments against using the client's sources only or following these to every detail, but this discussion, though important, is not central to this paper and shall not be continued herein.

What is important is that terms would be consistent within their environment, or at least – the translated text, to the same extent as they are consistent in the ST. There are plenty of examples of translated texts in which the translator has not used the same equivalent to one word throughout and the editor has failed to notice it. For example, a report by the European Migration Network, written in English and translated into Estonian includes two of such mistakes in subsequent sentences[1]:

> Example 2: "Uurimustöö eesmärk on analüüsida seost viisapoliitika ning rände juhtimise ja kontrolli, sealhulgas *ebaregulaarse rände* tõkestamise, vahel. Samuti informeerida poliitikakujundajaid ja analüütikuid viisapoliitika mõjust migratsiooni juhtimisele nii seadusliku rände soodustamise kui ka *ebaseadusliku rände* tõkestamise vaatepunktist." (Kaska 2011a: 5; italics and underlining by the author of the thesis)

The underlined term is 'migration management', which has been translated as 'rände juhtimine' as well as 'migratsiooni juhtimine'. An unknowing reader might get confused as to

[1] Admittedly, this text is a report, not a contract or a law and is by no means legally binding. However, considering that the client is an EU institution, and that these are known for having very strict demands for translations, this text should be consistent in terms at least within itself, if not the broader translation environment.

whether ‘ränne’ and ‘migratsioon’ refer to the same object. Checking from the original text (Kaska 2011b), they do. ‘Ränne’ is the Estonian equivalent of ‘migration’, preferred to the loan word ‘migratsioon’, which has also been widely used and is due to its similarity to the SL word, tempting to use. According to Aime Vettik (2011: part III), general principles of terminology state that the pairs of own and borrowed terms are permissible variants. However, she continues, the use of a loan word is only justified when the own word equivalent is not as precise or it is not as compact or generally known as the loan word that can easily be fitted in the sentence (ibid.). Knowing Estonian, this would justify the use of the word ‘migratsioon’, yet it still does not justify the seemingly arbitrary interchangeable use of both equivalents.

The other translation mistake in the paragraph (see Example 2) is the inconsistent translation of the English term ‘irregular migration’, where irregular is meant as ‘illegal’, since the context allows to draw the conclusion that what has been meant is ‘irregular’ as “not conforming to established rules, customs, etiquette, morality, etc.” (Dictionary.com, LLC 2012) rather than “without symmetry, even shape, formal arrangement, etc.” (ibid.) or “not characterized by any fixed principles, method, continuity, or rate” (ibid.). However, the fact that the English word ‘irregular’ has taken the additional meaning of ‘illegal’, does not mean that the Estonian equivalent of ‘irregular’ (i.e. ‘ebaregulaarne’) would have behaved in the same manner. Estonian grammar dictionary defines ‘ebaregulaarne’ through ‘irregulaarne’ meaning ‘unequal, inconsistent, random’ (the author’s free translation) (ÕS 2006). It can be concluded that the translation ‘ebaregulaarne’ is simply wrong. However, the problem covered by this thesis is that the translator has not made a difference, or if so, has not, after having changed her mind, gone back to the previous sentence to correct the mistake first made. This also enables the conclusion that the translator has not worked through the terms in the text

prior to translating which may have helped to avoid making these mistakes. The relevance of this issue is enforced further by the fact that the translation in question has been signed with the name of the translator as well as the editor. The author of this thesis is not familiar with either of them or their other work, and hereby wishes them no harm, but believing only experienced translators would be commissioned such a job, the example suggests that even translations that we as translators are willing to sign with our name are prone to such mistakes. This in turn proves how little attention translators pay, or are able to pay, to these important details during the translation process, be it due to time restraints or simply human imperfection.

This raises the question whether there might be something we can change in our translation methods or process to learn to "unmake" such mistakes. After all, even the best of us have room for improvement. Or, as Isaac Bashevis Singer has said, "There is no such thing as a good translator. The best translators make the worst mistakes. No matter how much I love them all, translators must be closely watched." (Goldblatt 2002)

The most important watch over translators is carried out by editors, who read through and correct texts after translation. However, as we could see from Example 2, this is not always fully effective. Editors are just as human as translators and therefore may also fail to notice such mistakes. This thesis continues to investigate whether it is possible to make certain changes in our used translation processes by changing the use of computer-assisted tools in a way that would help to avoid translating a term in two different ways within the translated text or its context.

1.1 HOW TO WORK WITH RECURRING TERMS

When it comes to working with recurring terms, there are a few important factors to pay attention to: (i) when do terms get their equivalents – before the actual translation or during it; (ii) whose memory is used for the memorising task – of the translator or of the software tool; (iii) when are the equivalents of terms applied to the TT?

Christiane Nord (in Saluäär 2007: para 8) has stated that theorists come together in believing that it is essential that translators should carry out a thorough text analysis before starting translation since that is the only way to ensure that they understand the text correctly. However, thorough analysis will aid the translation process in more ways than by just contributing to the translator's understanding. Nord's extensive model of analysis that incorporates 17 levels (Pym 1993: 4) includes also a section on vocabulary, but fails to take into account the need of vocabulary analysis that would enable term recognition and rate of occurrence. However, these are the factors that would help the translator ensure term consistency, since knowledge of recurrence would alert an attentive translator to check that the same equivalent would be used. While term recognition is very much an intuitive task for the translator, rate of occurrence could be determined using computerised tools, some of which are built for this purpose.

In a translation class at Tartu University in spring semester 2012 lecturer Ilmar Anvelt gave a group of students a list of terms to which they were to find equivalents. For the following class the same students were to translate a text which contained all the terms that they had previously found equivalents to. When the lecturer later asked for feedback about whether the students found it easier or more complicated to translate after the equivalents had already been found, there was no student in class saying that such a work method had hindered their work,

instead, as Anvelt later put it, “Students seemed to have found it helpful” (personal communication 29 April 2013). In private discussions with the author some students later explained that such ordering of tasks had made translating more comfortable because they did not have to cut the translation flow for term search and that they had been able to concentrate more on looking at the text as a whole. According to Baer and Bystova-McIntyre’s (2009: 159–183) beginning translators tend to focus on the level of the word or phrase, ignoring textual elements. In this case it can be inferred that the students were able to divide their focus so that the word- or phrase-based phase, including looking for equivalents, would be done first, enabling therefore the students to work with the other elements of the text – the macro-elements – in the TT formation (or “translation proper”) phase. Thereby, it should have been easier for the students to consider the text globally and therefore they should have created a TT that could be considered “of higher quality”. Actual testing on whether there were fewer translation mistakes and whether the translations were generally better, has not been made within this paper or by the lecturer in question, but deserves research. It is also not known whether any of the students followed the same scheme for their next translation.

There are different work methods translators can apply when working with recurring terms, which, as marked in paragraph 1 of this subchapter, can be divided by the stage in the translation flow in which they are dealt with. Also, Anvelt has noted that generally there are two options and that he divides translators into two groups – some who work linearly, solving all the problems as these crop up, and others who start from making a translation with as little help as possible to return to the then already translated text for changes and research (personal communication 29 April 2013). Carrying certain operations out prior to translation proper is known as “pre-translation” (Kenny 2011: 470). One of such pre-translation operations could

be inserting TL equivalents into the ST resulting in a text where certain segments (e.g., terms) have been pre-translated into the TL, whereas the rest remains in the SL awaiting translation (ibid.). Dorothy Kenny (ibid.) finds such pre-translation to be of "obvious benefit to those commissioning translations", but admits that others (referring to Julian Wallis 2006) have found that it is not always the preferred mode of operating among translators. Wallis (2006: 4) conducted a pilot survey to find out whether pre-translation improves translation quality, translator productivity and satisfaction or not, whereas in his pre-translation model translation memory system proposals were inserted automatically into a source text, producing a hybrid text containing a mixture of source and target language elements. Already in the early 1990s Claude Bédard (in Wallis 2006: 50) offered a system whereby terms would be changed in the TT with their TL equivalents creating such a bilingual, or hybrid, text. Whereas the created pre-translation text would simply be a SL text containing specialised TL terms and the actual syntactic structure of the ST would remain unchanged (Bédard in Wallis 2006: 442). Such replacements follow certain logic as they correspond to semantic units – nouns are replaced by nouns, verbs by verbs, etc., and therefore the hybrid text, despite being in two different languages that may be grammatically very different, should still be comprehensible for one that is fluent in both languages, i.e. the translator.

Another important aspect in working with recurring terms is the use of memory. We could distinguish between two kinds of memorising tasks: (i) the task to memorise the equivalent of a term that may recur in a text; and (ii) the task to memorise which terms have already appeared in the translated text. When in the first case a good 'tool' to be used to aid memorising would be a termbase, especially an ad hoc termbase, in the second case other kinds of methods can be applied. Whether the translator needs to carry out only one or both of

the memorising tasks depends on the chosen translation tool and method. The selected equivalent could be applied immediately throughout the whole text or separately every time when the term comes up in the course of translation or a tool could be used that gives the translator a warning about recurrence. If the translator does not give either of the tasks away he or she does not only have to pay full attention to which equivalent the term gets, but also that it would be applied consistently throughout the TT. Both of these memorising tasks can be aided by CAT tools and methods for that are plentiful.

The third aspect important in working with recurring terms is when the chosen equivalents of terms are applied to the TT. This coincides with the second memorising task and the possible work methods are is discussed in more detail in Chapter 3.

1.2 CAT TOOLS AND TERM TRANSLATION CONSISTENCY

The list of CAT (computer-assisted/aided translation) tools available to translators is extensive and it is hard to imagine a professional modern translator who would not use one or many of these in their daily work. Koby and Baer (2003: 211–227) stress on the importance of fully recognising that technology is of central importance in the translation activity in the 21st century. While it cannot be assessed how many people still use the paper-pencil approach, people who engage in commissioned professional translation (in which the consistent translation of terms is relevant in order to achieve client satisfaction), translators must become increasingly more computer literate and comfortable in using CAT tools. Especially since, as Lynne Bowker (2002: 6) explains, CAT tools are any and all computerised tools that translators use in their work. That may include anything from word processors and spreadsheet applications to translation memory programs, terminology extraction and management

software, and even corpus-processing tools, if used directly for translation purposes. Therefore, such tools form an integral part of the translation process and are developed and used to facilitate that process (Vandepitte 2008: part 3). Many of these tools can also help the translator in ensuring the consistent translation of recurring terms.

One of the most important features of CAT tools that has to be borne in mind is their difference from machine translation; i.e. the fact that CAT tools do not attempt to replace the human translator, but are rather used to facilitate the translation process (O'Hagan 2009: 48). It is still the translator who has total control over and full responsibility for the translated text and the translation process. That is one of the reasons why translators must have very good understanding of how such tools work and how these can be employed. This paper can only comment on how CAT tools can aid the translator in achieving the consistent translation of recurring terms and does not attempt to be in any way comprehensive.

Next in this chapter, a general instruction of three types of software tools – translation memory tools, terminology management tools, and corpus processing tools, is given with the aim of touching upon some of the main issues concerning each of the named types when trying to achieve consistency in terms.

1.2.1 Translation memory tools

One of the most commonly recognised CAT tools is translation memory software. Hatim and Munday (2004: 114) have stated these to be typical of the translator's work nowadays saying that "it is clear that the translation memory tool plays a key role in assisting translators." These are programs that work as databases that store ST and TT in segment pairs in a way that when the same or a similar new segment comes up later, the program offers a stored translation equivalent for reuse. Such matches are classified as exact matches, full matches and fuzzy

matches (Bowker in O'Hagan 2009: 48). But even in case of full matches, it is up to the translator to accept the offered match in full, change it, or reject it completely. To aid the translator in making that decision most translation memory programs (see list of commonly known providers in Appendix 1) display information on the comparative ST segments.

Translation memory tools are advertised as tools that enable the translators to save time as well as the trouble of having to translate one segment more than once. Whether and how well the programs are able to do that is irrelevant to this thesis. What is relevant, though, is the fact that not needing to retranslate a segment, but only the part thereof that has changed, is also said to ensure better consistency (SDL 2011b: 8). That opinion is supported by Hatim and Munday (2004: 113) who state that using translation memory programs "helps to assure consistency of terminology and means that translators only need to translate the changed text". However, Hatim and Munday seem to disregard that consistency is enabled only if the full segments are similar enough and would not operate quite the same way on the level of a word or a string of words. SDL (2011b: 8) also states that a translation memory program ensures that "no repeated occurrences will be translated differently", without acknowledging that translation memory programs compare full ST segments, and not phrases or words within these segments. According to Kenny (2011: 465) translation segments are sentences, headings, cells in tables, items in a bulleted list, etc., since these are easy to identify during automatic segmentation. Translation memory tools do not warn the translator of a recurring term unless it is surrounded by a similar amount of similar other text elements to a previously translated segment – therefore, e.g., if a term appears in a title and then in a longer sentence, it might not show up as a match. Some software providers have attempted to solve this by a feature called Subsegment matching (Bowker 2002: 103–5). Kenny explains:

> "The function works in a way that if no match is found in memory for a new ST segment, then the translation memory tool seeks matches for chunks of that ST segment. In order to do this the tool needs to have capabilities more commonly associated with example-based systems." (Kenny 2011: 469)

However, Kenny (2011: 464) also states that translation memory programs can be seen as quality-enhancing resources "in translation scenarios where consistency is at premium". This is most simply enabled in the segment-to-segment extent. Where the word or sting of words that requires consistent translation, i.e. a term, needs an equivalent to be retrieved from the translation memory, the program itself is not sufficient, but the translator's personal memory, and ability to hesitate, is to be drawn from. However, if the translator needs information on previously chosen equivalents, these can easily be retrieved using the concordance search function whereby the translator types in the search word and the program looks up all previous occurrences of that word in the translation memory. Although doubtlessly useful, a concordance search can help the translator only if the translator has a doubt about the equivalent. In the case of Example 1 (see Chapter 1) the translator could not feel it necessary to check the translation of such an easy word as 'ostja', but simply write down the equivalent to the word that he or she assumes to be correct. Therefore, a regular translation memory will only be able to take over one of the two memorising tasks.

Before concluding that translation memory tools are not as yet fully effective in ensuring the consistent translation of terms, it must be considered that many of these tools are provided with or recommended to be used together with termbases or other terminology management tools.

1.2.2 Terminology management tools

Termbases are similar to electronic dictionaries and generally contain single words and expressions, as opposed to general language dictionaries where whole segments of information

are stored about an entry word. Hence, tembases are narrower and can be used for more specific purposes. For example, when translating insurance conditions, an insurance termbase may prove more convenient than a general standard dictionary. Creating and maintaining termbases is especially useful for storing organization-specific, customer-specific or product-specific terms, which cannot be found in standard dictionaries (SDL 2011b). If the scope of a translation job is considered sufficient, the formation of a termbase for the specific purpose of the translation job, a so-called ad hoc termbase, may be drawn up before or during the translation process.

Terminology management tools help to do that. Even if the tools are to be downloaded and/or purchased separately, many of these can be accessed from within a translation environment, especially if both are provided by the same developer, e.g. SDL MultiTerm in SDL Trados Studio. The better these two types of programs are integrated, the more beneficial for the translator. For example, if the translation memory tool automatically searches for relevant terms from the termbase and warns the user of matches in some way that the user can rely on, the user no longer needs to doubt that every word in the text may be a term. Meaning, together the two programs can take up both memorising tasks and the translator is saved the trouble and time of having to look each word up in the ad hoc termbase. Such a “warning-system” in the translation memory environment as is added by a termbase program can help to avoid situations whereby not remembering that the term in question needed particular and consistent translating the translator uses the first equivalent that pops to mind or a new equivalent that has not been used in the text before. Admittedly, this integration function only works if work with terms has been done beforehand considering the specific needs of the translated text, or if the termbase is updated meticulously throughout the translation process. If

the used termbase has not been updated to the exact needs of the translation job, the user must still approach every word or phrase wondering whether it is a recurring one that might also need a consistent equivalent. Therefore, to gain the maximum benefit from using a termbase tool in assuring the consistent translation of terms a certain input is necessary from the translator.

1.2.3 Corpus-processing tools

Any kind of text available electronically can be a corpus and therefore, with appropriate tools, processed and analysed as such. These appropriate tools – corpus-processing tools – are computer programs enabling to access, manipulate and display data from a selected corpus. Using a corpus-processing tool one can retrieve such a complete word lists or keyword lists from the corpus and form ad hoc termbases thereof. Keyword-in-context features allow also the analysis of selected words, which aid the formation of such a termbase.

Storing and exploring corpora electronically using computerised storage and search facilities and concordancing programs has become increasingly common since the 1960s (Malmkjær 2005: 116). Much attention has been paid to corpora by translation theorists and student translators, less by practicing translators, whereas the use of corpora by all seems to be a relatively recent trend – Hatim and Munday (2004: 118) write that electronic corpora "are becoming increasingly used in research across the board in Translation Studies." However, it is not only theorists and researchers that can benefit from corpus processing, but practicing translators as well. Although Guy Aston (2009: ix) states that using corpus has not become widely established among professional translators, according to Kenny (2011: 462) there is evidence that "practitioners are making increasing use of monolingual target-language corpora (see Mahler, Waller, and Kerans 2008)." What Kenny (2011: 461–462) suggests is that

frequency-ranked or alphabetically ordered word or term lists and keyword lists can make starting points for the preparation of corpus-specific monolingual termbases in situations where terminology needs to be decided upon before translating. Naomi J. Sutcliffe de Moraes (2008: 26–33) is also concerned with using monolingual TL corpuses in preparation to translation.

Yet, as such, corpora can also be used in the SL for forming a ST-based termbase including terms relevant to the actual translation job. In Chapter 3 one of such concordancing tools, namely AntConc, has been selected as an example to display how to carry out such term management tasks – how to form ad hoc termbases based on monolingual ST corpuses.

Using an example given by Pilar Sánchez-Gijón (2009: 120), if the translated text includes the term ‘room temperature’ more than once and on one occasion it appears in the diagram of a control unit of an air conditioning system and in other occasions with hyponyms: ‘current room temperature’ and ‘desired room temperature’, the translator must pay attention that the equivalents chosen for all three include the precise term so that the TT user can recognise the connection. Now, if this information is retrieved from the text already prior translation, the equivalents can be decided upon and are ready to be used while actually translating the text. One way for retrieving that information is processing a ST corpus in a concordance tool that opens a keyword-in-context concordance for viewing (see subsection 3.1 AntConc). Already in 1995 Mona Baker, in her paper *Corpora and Translation Studies*, stated that the most important of the corpus tools that has potential use in translation is the keyword-in-context concordance, since it allows to display all instances of a particular word within the text in their surroundings and with their collocates (Hatim and Munday 2004: 118). Viewing terms in such a form (see Figure 3 on page 49) enables us to retrieve terms from the corpus with their co-

terms and variations, as in the ‘room temperature’ example. Also, keyword-in-context helps to find all terms used in the ST corpora. That is to say, the word ‘motor’ by itself may not be considered a term within the context of a car insurance contract, whereas viewing the word in the keyword-in-context list uncovers the word is commonly used within the corpora in pair with ‘vehicle’, and therefore the word pair becomes a term (‘motor vehicle’) that requires consistent translation throughout the text.

There may be various reasons why corpus-processing tools used for analysing the ST are not better known as CAT tools. One of these may be little knowledge of the option, since corpus-processing tools are rarely advertised for practical translation purposes. The other, connected reason may be that most articles talking about using corpora and corpus-processing tools for translation are more concerned with TT and TL than ST and SL, aiming therefore for wider research. ST-based term search is much more specific to the task in hand and translators, who are always pushed for time and cannot therefore educate themselves in the subject on a wider scope and create comprehensive TL corpora for the whole subject, will not engage in using corpus-processing tools. However, ST-based corpus is narrower, covering only the terms the translator actually needs in order to fulfil their particular job task. Aston (2009: ix) states that the “difficulty of using corpora is in that they rarely provide an immediate answer to a translator’s problem” and argues that corpus is not analysed by translators mainly because it is time-consuming. Whereas, as argued herein, that concerns TL corpora rather than specific ST-based corpora. Of course, even with ST-based corpora the data has to be analysed and evaluated and that takes time and effort, but less than in the case of TL-based corpora which may include much information that does not directly concern the particular translation job.

Another reason for the little use of corpus-processing tools may be of a more practical kind – translator's limited computer skills. Lack of statistics in this regard motivated to ask about computer study experiences of active translators in the thesis questionnaire (see Chapter 2). Either way, it can be assumed that few translators are introduced to a corpus-processing tool in an academic environment, and if they are, basing on anecdotal evidence, it is more likely that the tool is introduced more for text analysis than for straightforward translation purposes. Patricia Rodríguez Inés agrees that teaching such tools to trainee translators can improve their competence as translators:

> "If, as translator trainers, we wish to develop our students' competence to solve translation problems, then we need to provide them with strategies to use existing resources and tools, to create new ones and to reap the maximum benefit possible from them. /.../ We suggest that one of the learning objectives within a translation course is to grasp how to use corpora." (Rodríguez Inés 2009: 129)

Teaching translators to become independent users of corpora as part of their translation competence can improve their overall work results. Learning how to make best use of available tools and how to better analyse and process information should rank high on any translator's, trainee or not, list of interests.

2. HOW TERMS ACTUALLY GET TRANSLATED – STUDY

This chapter discusses the results of a study that was carried out in the form of a questionnaire in Estonia in February 2013. Its aim was to find out how different translators go about working with recurring terms from the aspects that would help to ensure the consistent translation of the same. The questionnaire was put up on eFormular.ee, a web tool for conducting surveys via the Internet. The link to the questionnaire was sent in an e-mail to two universities in Estonia teaching translation, 40 translation bureaus operating in Estonia and a small number of personal contacts. The letter specified that the respondent must either be a professional or a trainee translator. In order to avoid responses from people who have no actual contact with translating, all questionnaires where the respondent had not given information on their translation background, i.e. not replied to any questions from 1 to 3, were to be deleted (for the questionnaire see Appendices 2.1 (in English) and 2.2 (in Estonian)). No such responses were submitted.

In order to better accommodate the respondents and to draw a wider group of people, the questionnaire was intended to be language independent. Yet, to set some restrictions, it was made available in English and Estonian, whereas each respondent was free to choose their preferred language. For the purpose of analysis, answers in both languages were seen jointly without any differentiation.

The first part of the questionnaire was, as mentioned, intended for receiving information on whether the respondent has actual contact with translation – whereas the option “no experience” was omitted to discourage people with no translation experience from continuing – and to be able to group the respondents according to their experience to receive more

information about the correlation of study background and translation experience with work methods.

Since in part this paper stresses on the importance of computer literacy and teaching computer skills to translators, the respondents were asked whether they would like to improve their relevant competences. Although the question presumes a yes-or-no answer, an open box was provided to encourage the respondents to reason their choice, without forcing them to do so – i.e., the box could also be used for marking “yes” or “no”. The expectation was that people that do not wish to provide a more explicit answer would simply give the answer needed for the statistics, whereas others would help to explain why the answer inclines towards one or the other end. Since this paper is not directly concerned with translator training or computer study need, more specific information on that matter did not seem relevant at the moment of compiling the questionnaire.

The main objective of the questionnaire was to find out more about the respondents’ translation habits that are relevant to the consistent translation of terms. For that purpose, the respondents were asked whether they take time to work through the text looking for terms before actually starting to translate or if they tackle every term separately as these come up in the course of translation. To get more information about work methods, the respondents were also asked if they read the text through before translating or not. The answers were broken apart to three alternatives: reading carefully, skimming, and not reading. This was to avoid a situation where people that only scroll through the text reading it here-and-there would be tempted to answer “yes”. The question was followed by two additional questions: whether the respondents mark terms in the text when reading it; and, if they do not read, whether they find terms from the text using any kind of software.

To ensure more consistent and better comparable responses, a translation case was made up which the respondents were asked to keep in mind while answering the questions. Also, the assumption was made that asking the respondents to focus on a particular translation job gives an opportunity to keep them from feeling drawn between replies. Although Tavast (2008: 96) states in his survey that "it appears that translation habits are generally not changed during one's career", that does not necessarily mean there are no variances to how a translator approaches a text depending on, e.g. its scope, the deadline and the text itself (subject matter). To avoid a situation where respondents would like to answer "it depends", they were guided to a particular situation where the choice of actual translation method would not "depend" but be chosen according to what is known. In describing the translation case, it was made sure that the ST would be of a kind where the inconsistent translation of terms would be considered a translation mistake and where there would be an obvious text environment. The described translation task is a 17-page translation order consisting of two MS Word documents. One of these documents is a contract, the other, general contract conditions annexed to the contract. The translator is given three full days before having to submit the finished translation. That leaves a little over five pages per day to translate, which can be considered relatively generous, so the translators should not feel pressed for time.

Next the respondents were asked which CAT tools they would choose for translating such a text. Here, possible options were offered. The last question was concerned with the translator's behaviour attempting to receive information on how much translators rely on their personal memory in making sure terms get the same equivalent on each occurrence in the text.

Due to the need to remain universal, it seemed justified to use simple comparative interpretation of the received data. Also, weaknesses in the study results must be considered. It

can be assumed that the respondents might have thought the aim of the questionnaire was to prove that they do not use available tools to the best of the abilities of these, and may have wanted to argue with that somehow. The letter sent to the possible respondents as well as the questionnaire itself did not include an explanation of the aim of the questionnaire or a brief about the exact nature of the study. On the contrary to the intention, this may have created some confusion instead as some respondents took the chance to lecture on how they find machine translation being “useless” and how no computer program or artificial memory is in their opinion capable of replacing the human ability to think or memorise. Also, a few respondents considered a 17-page translation being too short to bother setting up a termbase or marking terms in any way, one even considered the scope insufficient for using a translation memory tool. However, to make using a translation memory program even more reasoned, the translation task had been set up so that the text would be in two different files – increasing therefore the likelihood that there are recurring text segments.

2.1 STUDY RESULTS AND DISCUSSION

The study sample consisted of 89 people who filled in the questionnaire. 72 of them (81%) said to have studied or to be studying translation academically, which may refer to an overrepresentation of students. However, since the respondents said that their translation experience is mostly for work (58%) or only for work (16%), whereas only one respondent claimed to have translation experience only in connection with their studies, the ratio of actually working translators is high. Also, 42% of the respondents said their translation experience is extensive and 38% rated it to be medium. This may give reason to suspect that the line between medium and extensive, as specified in the questionnaire, might not have been well chosen and five years may not be an adequate assessment of long-term in the field. Still,

knowing that most respondents have relevant work experience and are mostly translating for work, the data can be used for drawing some conclusions on the issues concerned.

To begin with, the respondents were asked to give information about their experience with CAT tools[2]. Not to restrict the respondents and because such experience may differ considerably, an open box was left for specifications in free form. The respondents can be divided in two: (i) those who said that they have studied computer programs in some academic setting, either in school, at university, or at a training course (59%), whereas some of them (15%) chose more than one of the named options; and (ii) those who said that they have not studied any computer programs in such a setting (41%). So the number of self-taught computer-users is relatively high. Although self-studies may often be very fruitful, these are perhaps most efficient when the person knows which particular function they need to learn or which problem they need to tackle. When teaching oneself, a person is less likely to come across such functions of the program he or she was previously unaware of but that they might find helpful in their work. One of the respondents wrote, "I've studied [translation memory programs] Trados and WrodFast on my own. These are easy programs that can be learned in half a day. They all follow the same logic that one just has to get." Another respondent confirmed the opinion, "Programs are relatively easy. I use SDL products." Frankly, this kind of confidence is admirable. While it is understandable that certain kind of knowledge does come easier to some people, it is hard to believe that a translation memory program can really

[2] This statistics may be somewhat inaccurate since it seems not all respondents are aware of the fact that word processors or even an e-mail environment can also be considered a computer-assisted translation tool. To avoid confusion "CAT tool" was replaced by "computer programs you use for translating", but the respondents' comments give reason to believe that some may still have taken these for translation memory programs, in one case even for machine translation software.

be acquired in depth in half a day. For example, SDL offers three workbooks of different levels, each around 100 pages, for their translation memory program.

When asked about their willingness or necessity to improve their computer skills, (i) 72% of the respondents said to be interested in that, and (ii) 28% said this is something they are not concerned with. Almost half of the respondents in this second group (48%) had not studied computer programs academically, which points to a group of people who have not and do not want to be taught computer programs but are happy with teaching themselves or with no teaching at all. However, this is a minority, as almost three quarters of the respondents stated that they would like to study computer programs. Some of those people had studied computer programs before in one or another form and admitted that these studies, although useful, made them realise how many functions programs have that they are unable to use without further guidance and therefore they would like to improve their skills and knowledge. One of the respondents wrote, “[I’ve taken] one course as an exchange student at a foreign university. I’m guessing they were only able to explain us the very basics in that semester.” Another confirmed the opinion, “[I’ve studied] Trados briefly at university. I would be able to use it, but I believe there is much more to learn of the program.”

When looking at willingness to study and the respondents’ CAT tool study background together (see Chart 1), it is apparent that those respondents who have studied computer programs before are much more willing to continue such studies, whereas there is a remarkably higher number of people not interested in academic computer program studies in the group of people who have no such previous experience.

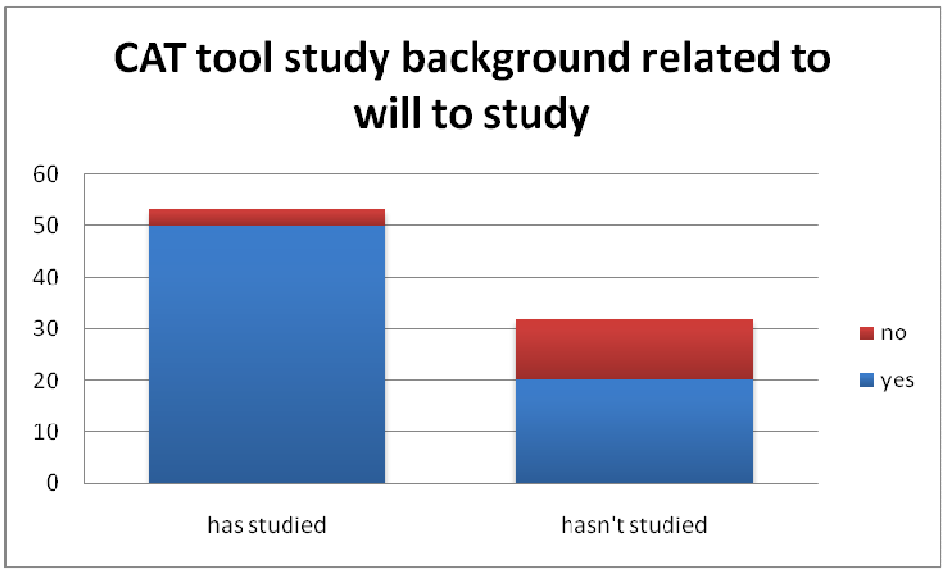


Chart 1[3]: CAT tool study background related to willingness to improve computer skills[4] (where "no" refers to "does not want to study CAT tools" and "yes" refers to "does want to study CAT tools")

Although opinions on whether studying CAT tools in an academic setting is fruitful and necessary or not, can be very different, there is a clear tendency among the respondents to study CAT tools and improve their relevant skills further, which allows the conclusion that this must be something translators find important in their job.

In the main part of the questionnaire, questions were set up about a translation case. To begin with, the respondents were asked which type of CAT tool they would use for translating the text described in the translation case. The responses divided as follows: (i) using translation memory software (e.g. SDL Trados, WordFast) (47%); (ii) in a word processor (MS Word) writing on top of the SL text replacing it (30%); (iii) using translation memory software with terminology software (e.g. SDL Trados Studio with SDL MultiTerm) (11%);

[3] Here and on all other charts in this chapter the number on the vertical axis refers to number of respondents.

[4] Here, a certain margin of error in interpreting the result deriving from the respondents understanding of the question must be admitted.

(iv) in a word processor (MS Word) opening a new file for the translation (11%). No respondents specified using any other kind of program and all respondents replied to this question, allowing the conclusion that all respondents use CAT tools in their work. Interestingly, there were also a couple of people who commented on being principally against using such tools. One respondent claimed not using translation memory programs because "the human ability of thought and memory cannot be replaced by any computer program or artificial memory". The author of this thesis is of the opinion that while human ability of thought should by no means be replaced by a computer in the translation process, human memory on the other hand is not completely trustworthy, and can be supported by a program that is more systematic. Although machines might need the human interference and need to be checked, such tools can aid us in assuring we achieve good results. In conclusion, the solution the author of this thesis would stand by is the combination of the best of two – human ability of thought and the ability of the computer to memorise – i.e., save, withdraw and display information.

The choice of CAT tool did not seem to be related to whether people have studied translation in an academic environment or not as the division was relatively even. But a link can be made between chosen programs and work experience: people with extensive work experience did not choose the option of opening a new file in a word processor, whereas people with short work experience did not choose using a translation memory program together with a terminology management program (see Chart 2).

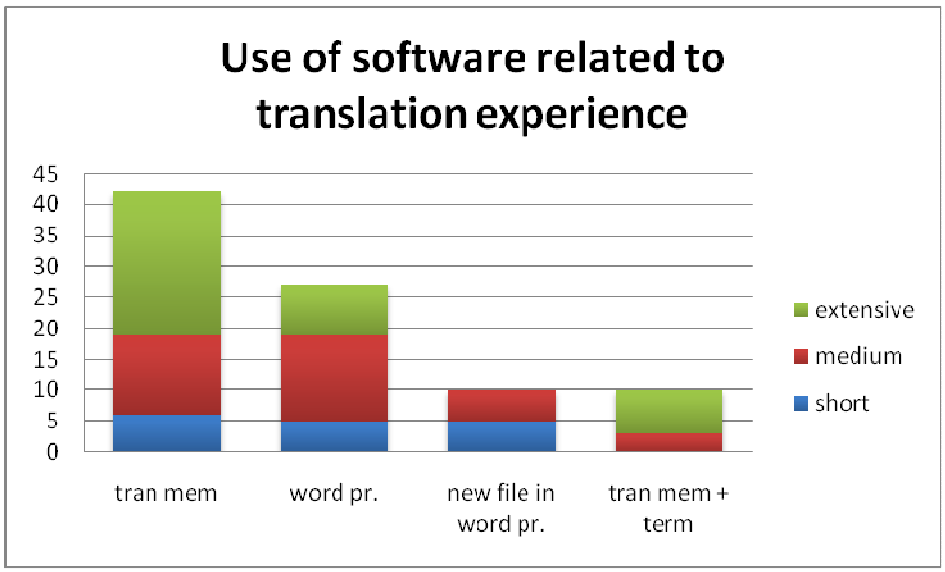


Chart 2: Choice of CAT tool related to translation experience

Overall, the respondents' choice of translation tool did not seem to have been influenced by their previous study experience. However, concerning self-taught respondents, it was noted that fewer of them chose opening a new file in a word processor for translating. Also, of the respondents that chose to use translation memory together with a terminology program, half had not studied programs in an academic setting and did not want to. Therefore, it seems, self-taught respondents do generally use more complicated programs in their work and tend to make a more computer-reliant choice. Interestingly enough the choice of translation memory with a terminology program was only made by those who had not studied CAT tools in an academic setting and did not want to (the so-called self-taught respondents), or by those that had studied some programs before and wanted to continue such studies.

The respondents were also asked to explain why they translate using the particular program they had selected. People who chose to use a translation memory program named the following reasons: previous memory can be useful; ease of work; habit; the ability to use concordance search; it is comfortable; memory can prove useful in the future; consistency;

employer's request; comfortable screen division solution; no threat that a sentence would be missed and not translated, etc. Some of these respondents also explained, that they use translation memory only because they have no access to terminology programs. Since the programs are to be purchased separately even if provided by the same developer, the additional cost may restrict the translator's ability to use these programs. Secondly, as a couple of respondents pointed out, the choice may be influenced by lack of knowledge of or experience in using such programs. But it also seems there are translators who simply do not consider it necessary. Many of the people who had chosen to translate using only translation memory software said that it is sufficient to search for recurring terms from translation memory (using the concordancing tool) or that when it comes to terms and their equivalents their own personal memory is sufficient. One respondent marked that the translation task (i.e. 17 pages) is too short to worry about setting up a troublesome termbase. A few others also noted that termbase creation takes too much time. Another respondent claimed using translation memory only due to "laziness" – supposedly that person is also too lazy to use termbase software.

The people that responded to use translation memory with terminology software gave different reasons for their choice: many claimed it to be the most comfortable way; some noted that it is the most elaborate, complete and therefore the best solution; one marked that usage is required by the client.

Those that chose to use a word processor reasoned their choice most frequently by saying that they have no access to translation memory programs, or that they do not know how to use these: Many respondents claimed that using word processors is easier and more convenient, one even claiming it is "less time-consuming and safer" than using a translation memory

program; one respondent uses a word processor because there is no requirement from the client to use anything else; two respondents said that 17 pages is not long enough to bother with opening the files in any other program but the word processor (the translation tasks specified the source text was sent as a MS Word file).

Therefore, reasons for choosing a certain program are different but all seem to come down to what is available to the translator, what is the translator's attitude towards the options and which benefits they believe the programs can offer. Also, it seems to be that beginning translators are less likely to be provided with programs by employers or are less likely to be asked to use programs by their clients.

In addition to the main program used for the translation, the respondents were also asked if they use any other programs to aid their work. Most respondents found aids being necessary in the translation process with 78% of all respondents marking to use aids amongst the offered options: SDL MultiTerm or similar termbase; AntConc or similar concordance program; self-made termbase for that particular translation (in MS Word, MS Excel or other electronic means); self-made termbase for that particular translation (on paper). Some of those respondents (22%) chose more than one option, which on the one hand points to their willingness and necessity to be helped in the memorising process, but on the other hand can also refer to some multiplication of tasks (e.g. writing the equivalent of a word down on paper as well as entering it into a digital termbase). However, when asked whether the use of aids is connected to finding terms in text as little as eight people, i.e. 9% of all respondents said they use some kind of software tool for that particular purpose. Three of these respondents chose to comment, marking that they find terms in the ST using AntConc (one person), ApSIC (one person), or SDL Trados together with SDL MultiTerm (one person). Therefore, while all

translators use CAT tools in their work, only a minority of translators use help specifically for finding recurring terms in the ST.

Admittedly, there are also other methods for finding recurring terms, e.g. the translator can go through the text looking for recurring terms without the help of software tools. As discussed in Chapter 1, translators' work methods differ also by whether they search for terms (and their equivalents) before translation proper or during it. That is why the respondents were asked whether they take time to read the text before commencing with translating. The responses divided as follows: (i) 64% said they only skim the text through, (ii) 26% said they only have a quick look and start work immediately, (iii) the remaining 10% said they read the text through carefully. Furthermore, only 5% or the respondents said that they mark recurring terms in the text in some way (underlining, highlighting, writing out, etc.). Again, only a marginal number of people engage in finding recurring terms in the ST.

The respondents' comments to the questionnaire conveyed the opinion that previous reading is considered unnecessary. One respondent wrote, "The previous reading of the text (including the previous reading of the sentence to be translated) is too much of a waste of time. If it turns out during the translation process that the thought takes a different direction or the initial translation (i.e. the start of the sentence) needs to be reworded, this would take less time than reading the whole text." Many others also noted that they find later editing to be more convenient as it is easier to pay attention to such things after translation proper. A couple of respondents admitted jotting down information concerning things that need to be checked after the translation has been completed to aid their memory in the editing process.

But, as discussed already in Chapter 1, memory has also an important role in the translation of recurring terms. On the basis of their answers, the respondents can be divided into three

groups: (i) a group that is totally dependent on their own memory in making sure terms get the same equivalent, both in terms of remembering the equivalent as well as which words must be treated as terms (i.e. recur) (23%); (ii) a group that does mark down the equivalent, but still relies on personal memory in remembering to check the equivalent of the word (50%); and (iii) a group that give the whole memorising task over to the computer (27%). The results show that nearly three quarters (73%) rely on their own memory in one or both of the memorising tasks.

It was also considered interesting to see whether people who are interested in improving their computer skills are more likely to be the ones who rely more on their own memory. But on the contrary, the results show that there is a considerably higher ratio of those who are not interested in improving their CAT tools skills in an academic setting amongst those who take up both memorising tasks (see Chart 3). Does this refer to a group of people who are independent of artificial memory and wish to remain so? The overall percentage of these people, however, is only 15%.

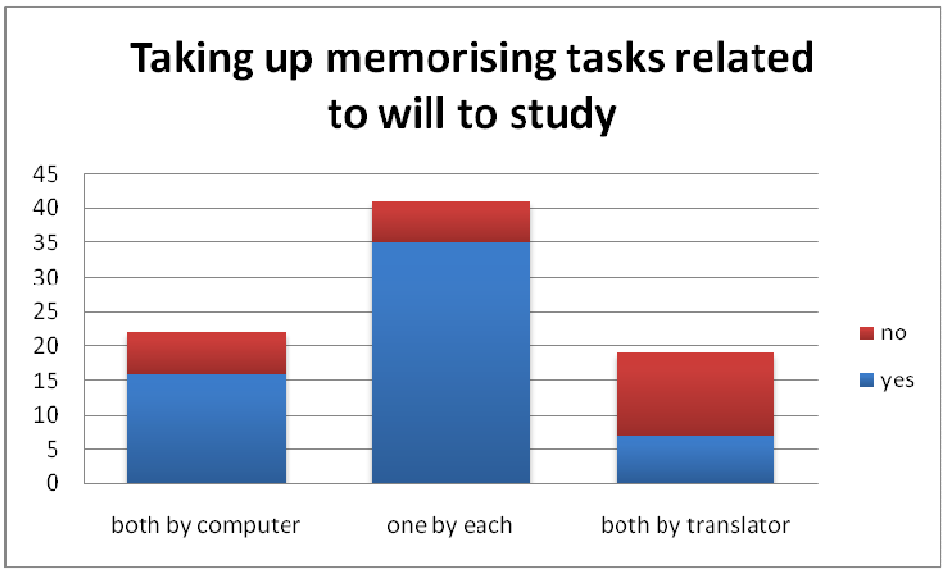


Chart 3: Taking up memorising tasks related to willingness to improve computer skills (where where "no" refers to "does not want to study CAT tools" and "yes" refers to "does want to study CAT tools")

The last question of the questionnaire illustrates how much translators rely on their personal memory. The respondents were asked what they do with a term after they have come across one and spent more than 5 minutes to find an equivalent for it. The responses divide as follows: (i) translate it in the one place and memorise the answer (20%); (ii) translate it in the one place and write the term and its equivalent down for yourself (e.g. in a word processor, a spreadsheet application or on paper) (55%); (iii) translate it and apply the translation throughout the text (e.g. using the Replace function) (10%); (iv) translate it and add it to a termbase (e.g. SDL MultiTerm) (15%). This confirms again that only a small number of respondents give memorising tasks away to the computer. The fact that a fifth (20%) rely only on their own memory in case of a 17-page insurance document, is alarming and points to the need to raise translator awareness on the issue of recurring terms. However, a fourth of respondents (10%+15%) can be considered to pay attention to such issues as their choice of

work method would allow the computer to help the translator by taking up both memorising tasks.

The aim of the survey was to find out more about translators' work processes and habits in connection to treating recurring terms. In order to get a better view of actual work methods, each respondent's answers were assessed separately, paying attention to the treatment of recurring terms as well as how the translator takes advantage of the tools they use. For the benefit of drawing conclusions, the respondents were divided into three groups: (i) finds consistency of terms to be important and relies on technical means in the memorising tasks (26%); (ii) uses technical aids to some extent, gives away one of the memorising tasks, pays attention to terminology issues (52%); (iii) relies on personal memory and uses technical aids as little as possible, often being principally against using CAT tools, does not mention considering term consistency issues (22%) (see Chart 4).

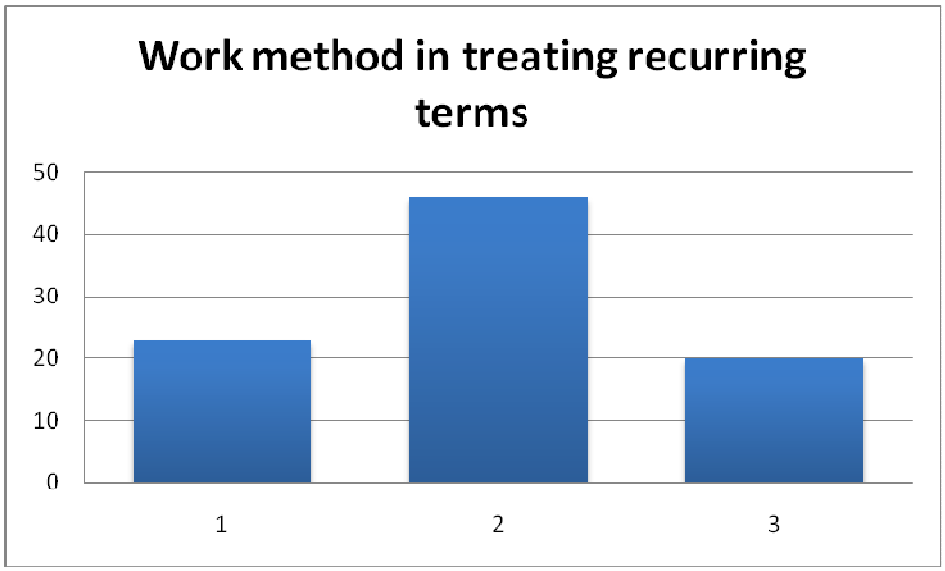


Chart 4: Assessment of the respondents' work methods when treating recurring terms

As can be seen, the number of respondents whose work method was assessed to belong to group 3, was the lowest. The work methods and comments of just more than half of the

respondents (group 2) lead to believe they are somewhat concerned with recurring terms getting a consistent equivalent, whereas the work methods of just over a quarter of respondents (column 1) reflect reliance on computerised tools and concern in providing a terminologically consistent translation.

2.2 CONCLUSION TO STUDY

In the preceding subchapter, the translators' habits were assessed by which tool they prefer to use, when they treat recurring terms, and whose memory they rely on. If to create an "average translator" based on the most frequent answers, the person would: not read the text through prior translation; use a translation memory program; rely on their own memory and the concordance feature of the translation memory program; treat terms as they come up and not apply the equivalents of these throughout the text. The study also showed there is little correlation between using computer programs in translating and having studied these in an academic environment. However, a large number of respondents wished to improve their knowledge in using CAT tools, irrespective of their previous work or study experience with the same. This shows that CAT tool skills are considered important for the work of a professional translator and that a good command of these is valued.

The study was able to meet its aim of retrieving information about how translators go about working with recurring terms from the aspect of ensuring consistent translation of the same. Based on the replies and the discussion in the previous subchapter, a few generalisations can be made. The most important of these are as follows:

- All translators use CAT tools to some extent and find it important to be able to use these
- The choice of work method based on various factors, most important of which are the availability of programs and the translator's attitude to using CAT tools

- More than half of the respondents have studied some or several CAT tools, whereas nearly three fourths wish to study more
- Translators trust their own memories
- Even translators who use technical means that would allow giving both memorising tasks to the computer do not often choose to do so
- A marginal number of translators look for recurring terms in the text
- Translators are more concerned with speed and ease of work than providing a term-consistent translation

The fact that even if the translators take the time and trouble to read the text through before translating (which very few admitted doing, even fewer mark recurring terms) and only a marginal number of translators apply an equivalent of a term throughout the text when first approaching the term enables the conclusion that translators rarely employ available technical means to make sure recurring terms get the same equivalent. Most translators whose answers and comments revealed they are concerned with terms getting consistent equivalents either trust their own memory or prefer to check these equivalents later in the editing-process. One respondent wrote, “When proofreading the text, I check with Replace, whether I have used the term everywhere where it is existent in the original text.” Another explained, “I check repetition in text with ctrl+F. Sometimes I have checked after completing the translating whether a term or a name is without typos everywhere it appears.”

The comments also revealed that translators are often more concerned with the speed of translating and how comfortable their work is. Also, that recurring terms are not a consideration for them. Some translators were also principally against term-work, one respondent writing, “I’m very sceptic about word-based translating, i.e. I don’t think you could translate basing solely on a dictionary/termbase, without understanding the content. I find content to be primary since the client will understand very fast if only words have been translated without understanding the content.” This paper is not advocating that kind of term-

based translation. Rather the opposite – giving these more "mechanical" tasks over to the computer would enable the translator spend more time on and pay more attention to the content of the text (the "understanding") and on how to build a more TL-like text.

The study raised many issues that would deserve more consideration. Some suggestions for further studies:

- Do translators who give both memorising tasks to the computer achieve better results in the consistent translation of recurring terms than those translators that rely on their own memory?
- A more comprehensive and structured view on why translators choose to use or not use particular types of CAT tools. On the one hand this would help developers understand the fears and expectations of their market, on the other it would explain the choice of translation methods and what translators find important.
- Comparing translators' attitudes towards using translation memory and terminology tools with their background in studying these.
- Research covering the specific computer literacy needs of translation students. The study group could include prospective, current, as well as former translation students. The study could touch upon teaching various CAT tools, as well as, e.g., typescript, within translation study programmes and the students' expectations as well as their satisfaction level. It could be carried out to receive more information about the needs of future translators and how the universities could be better equipped to cater for these.

3. SOFTWARE DEMONSTRATIONS

The programs used for these software demonstrations were chosen so that these would cover various types of software programs used by translators in their work. Since the list of such programs is extensive (see Appendix 1), the ones of each type that are most commonly used and easily available were chosen as test tools. Also, a concordance was added by the author to display a possibility for retrieving terms from a text before "translation proper". The selection included: (i) a concordancer (AntConc); (ii) a terminology management program (SDL MultiTerm); (iii) a translation memory program (SDL Trados Studio 2011); (iv) a word processor (MS Word).

In order to better demonstrate the software, a text was chosen for test translation. This text was to be both subject matter specific and client-specific, and was chosen from among the documents of a particular insurance company. The ST was in Estonian, approximately 17 translation pages in length (where one translation page is 1800 characters with spaces), and comprised of general motor vehicle insurance conditions for year 2012. The version for year 2011 of the same conditions, both in Estonian and English, was also available from the same source – the web page of the insurance company. In this chapter the named translation is not completed in its entirety, but only used for showing how various approaches to translating recurring terms may be performed with maximum success in achieving the consistent translation of the same. The aim was to show which means the translator can use to rely as little as possible on their own personal memory and whether and how the tasks of memorising can be left to the machine. To begin with, the text was processed using a concordance tool since that was able to provide an input for a termbase. In the next step, an ad hoc termbase was created using a terminology management tool. The work continues in a translation memory

tool. Also, the same translation is used as an example for demonstrating the relevant capabilities of a word processor.

3.1 ANTCONC

A concordancer is a tool that gives information on the frequencies of occurrence of words in a particular text and can therefore be used as a method for receiving information on recurring terms in a text to be translated without having to read the text through and trying to remember which terms have already appeared, or having to guess which are likely to appear again. One of the programs that enables such exploration is AntConc, which is a freeware tool developed and provided by Laurence Anthony, at Waseda University, Japan. In corpus linguistics this type of tool is called a concordancer and it imports user-selected reference texts from their original files to the program where it can show the contexts for all occurrences of a search term, displaying also its collocations (Sutcliffe de Moraes 2008: 26). AntConc is available from Laurence Anthony's web page www.antlab.sci.waseda.ac.jp.

Enabling AntConc takes only a couple of minutes. One may also choose to simply run, and not save, the program. In this case space would be left unused on the hard drive of the computer. However, this option is only available for use while connected to the Internet.

For opening files in AntConc these must first be saved as Plain Text (format: .txt). Since the test translation was provided as a docx-file, one must start with the conversion. The file can then be opened in the program (**File** -> **Open File(s)**, or Ctrl+F). The name of the opened file is displayed in the left-hand column **Corpus Files** of the program window (see Figure 1). More than one file could also be opened at the same time, which is useful if the translation job consists of more than one file.

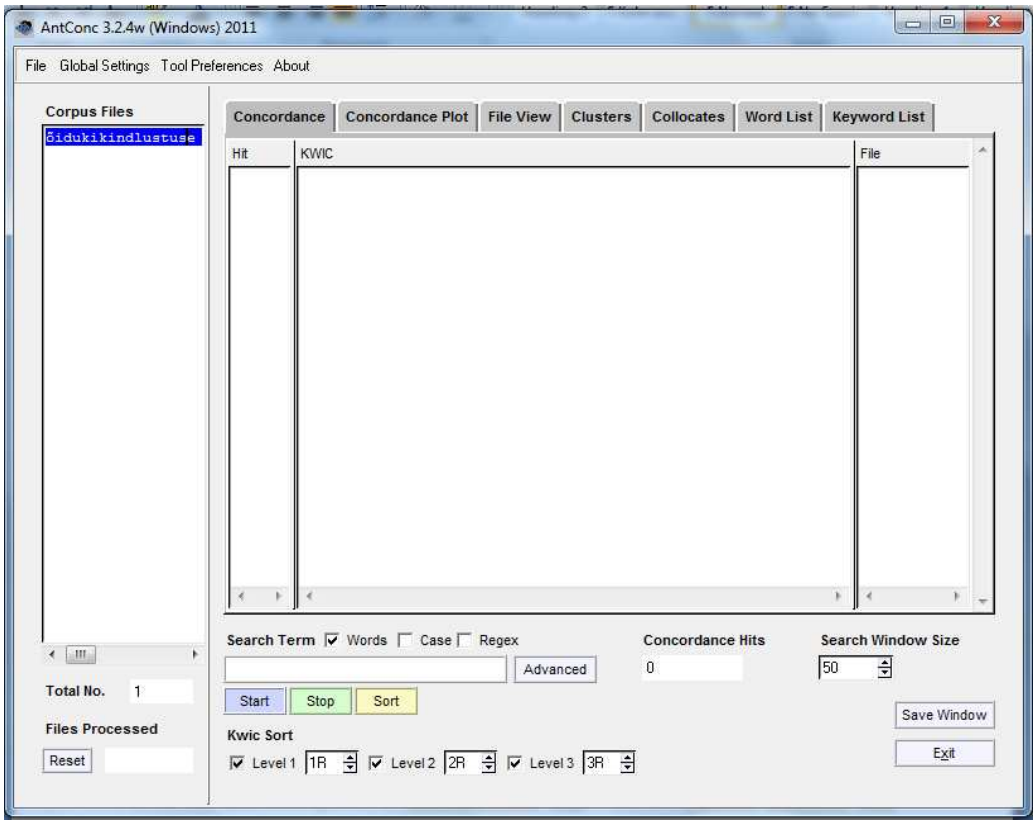


Figure 1: Creating corpus files in AntConc

To retrieve a word list by frequency, page **Word List** is selected (see (1) on Figure 2), search term is specified as "words" (see (2) on Figure 2), "sort by frequency" is selected (see (3) on Figure 2); since there is no need to separate uppercase and lowercase words, display option "treat all data as lowercase" is also selected (see (1) on Figure 2). To retrieve the information, **Start** button is clicked on (see (5) on Figure 2).

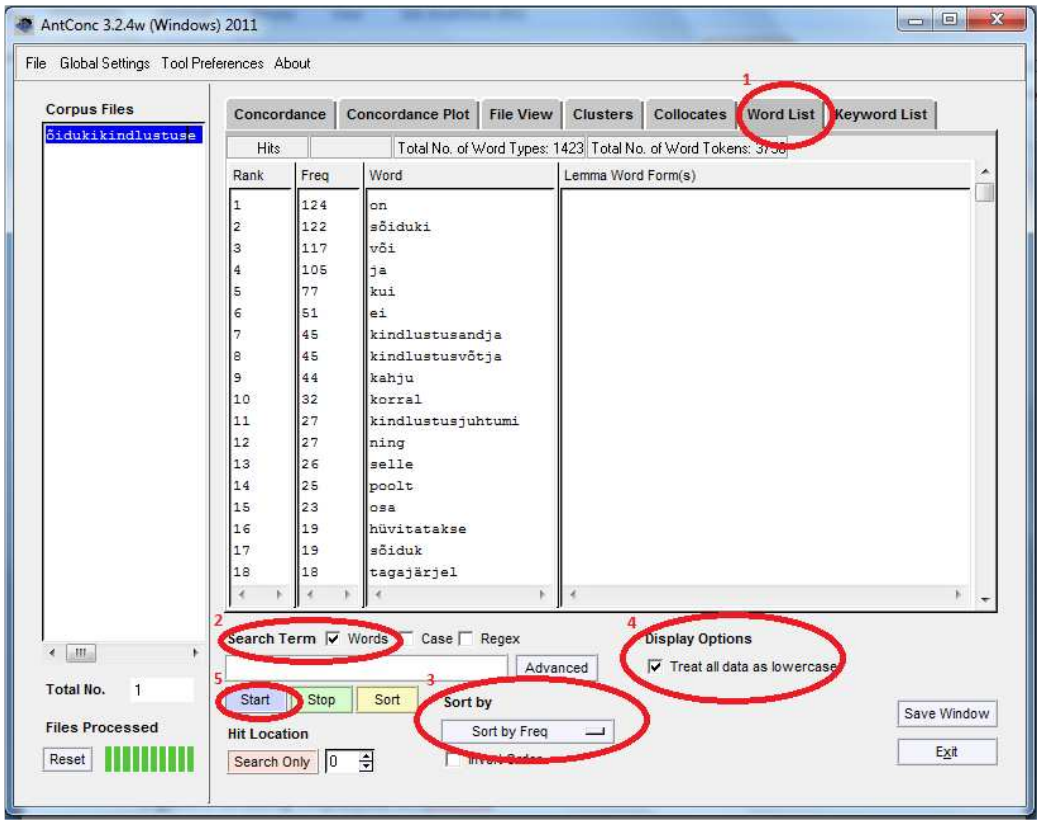


Figure 2: Retrieving a Word List by frequency in AntConc

As can be seen from Figure 2, the result appears in four columns: rank (in this corpus total of 1423 entries), frequency (most common word appearing 124 times), word (most common word 'on' (ET, verb, meaning: 'is')), and lemma word forms (none displayed). Since the word 'on' is a simple verb, it does not need to be treated as a term and can be skipped. The following word, however, can be considered a content-word and should therefore be checked in the concordancer. Clicking on the word 'sõiduki' (with 122 concordance hits) a new view (page **Concordance**) is displayed, where the word is shown in its keyword-in-context view, showing 50 characters before and after the word (the size can be changed under **Search Window Size**) (see Figure 3). This enables us to learn more about how the word is used in context and which words it is commonly used with, which helps to decide on a suitable equivalent.

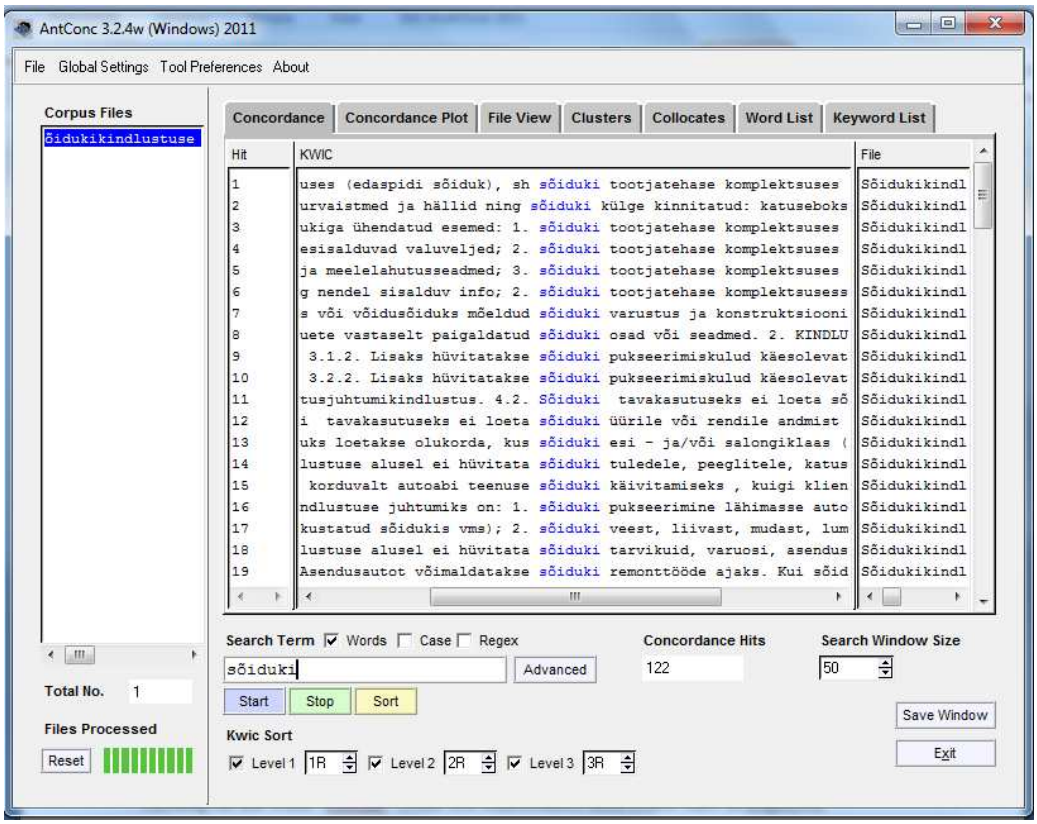


Figure 3: Keyword-in-context concordance view in AntConc

The concordancer gives the user maximum information on how the word is used in the particular text, all on one screen in only a few clicks. Without seeing the term in such context, the translator may unassumingly choose the wrong equivalent and find only later that he or she has made the wrong conclusion. If this happens before starting to translate, the mistake can easily be remedied but knowing that most translators do not do term-work before starting to translate, the same term can easily receive two equivalents in the same text. Working before translating with a concordancer will help to make sure that the final equivalent is chosen before equivalents are entered into the TT. Also, that inconsistencies caused by the translator not remembering that the word has already appeared and received a different equivalent could be avoided.

For getting a better overview of which words the search-word is commonly used with, it is advisable to hit the **Sort** button. In this case the display changes and the words following the search word appear in alphabetical order (see Figure 4).

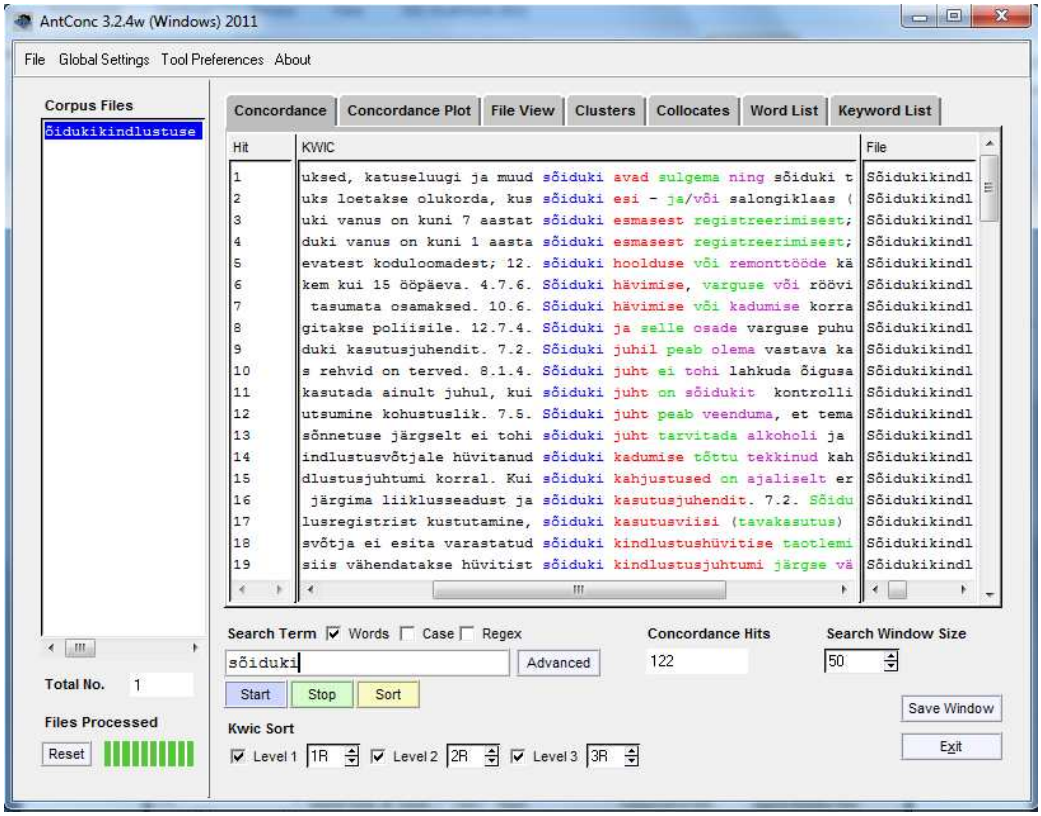


Figure 4: Sorted keyword-in-context concordance view in AntConc

By doing so, strings of words could immediately be treated as terms of their own and receive a common translation equivalent.

There is, although, a definite flaw in this search. As can be seen, the search word is not in its nominative form (the keyword is ‘sõiduki’, i.e. genitive). When going back to the word list and entering search term ‘sõiduk’, the nominative form, we see that there are 19 instances of that version of the word also in the corpus (see Figure 5).

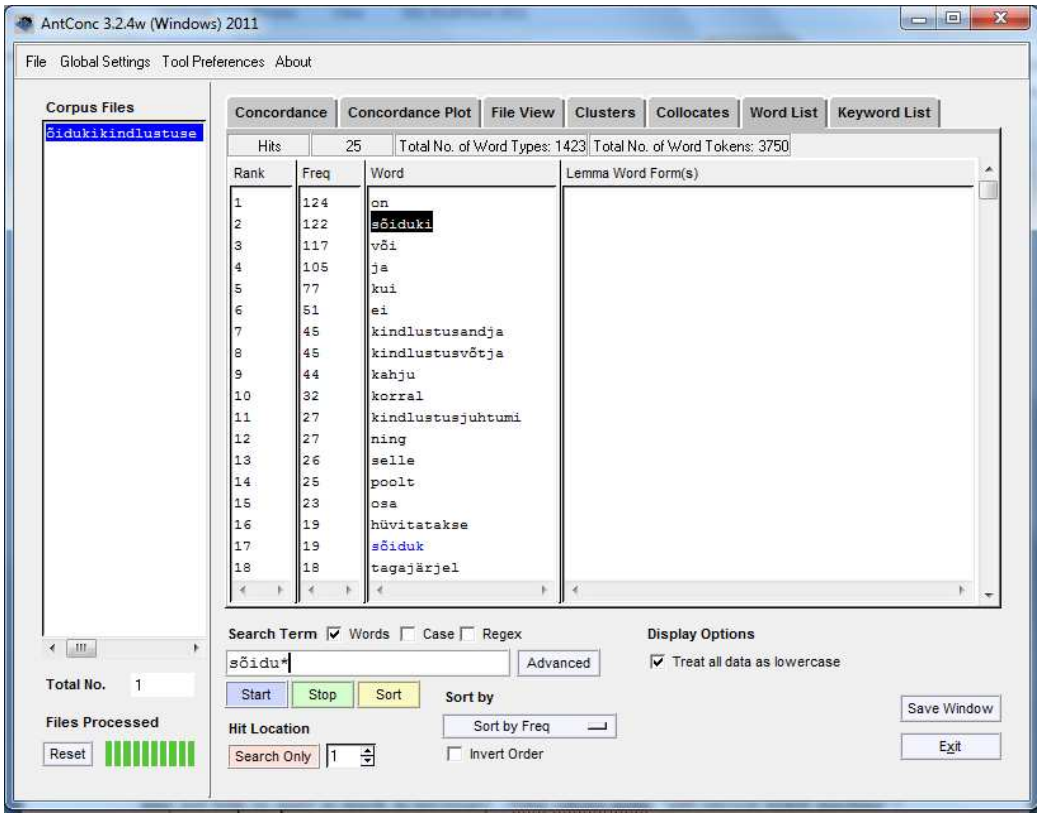


Figure 5: Concordance list displaying word forms of one word ('sõiduki' (122) and 'sõiduk'(19)) as separate entries in AntConc

To view all instances of the word in the keyword-in-context list we should enter the beginning of the word in the **Search Term** box with an asterisk where we have left letters out – 'sõidu*' (usually the stem of the word is sufficient). Now the program colours all words starting with these letters in blue and displays these in seemingly random order (as they appear in the original text) (see Figure 6). Note also that now the list includes all together 192 concordance hits, meaning there were more hidden forms in the corpus than the two that we were able to detect at first.

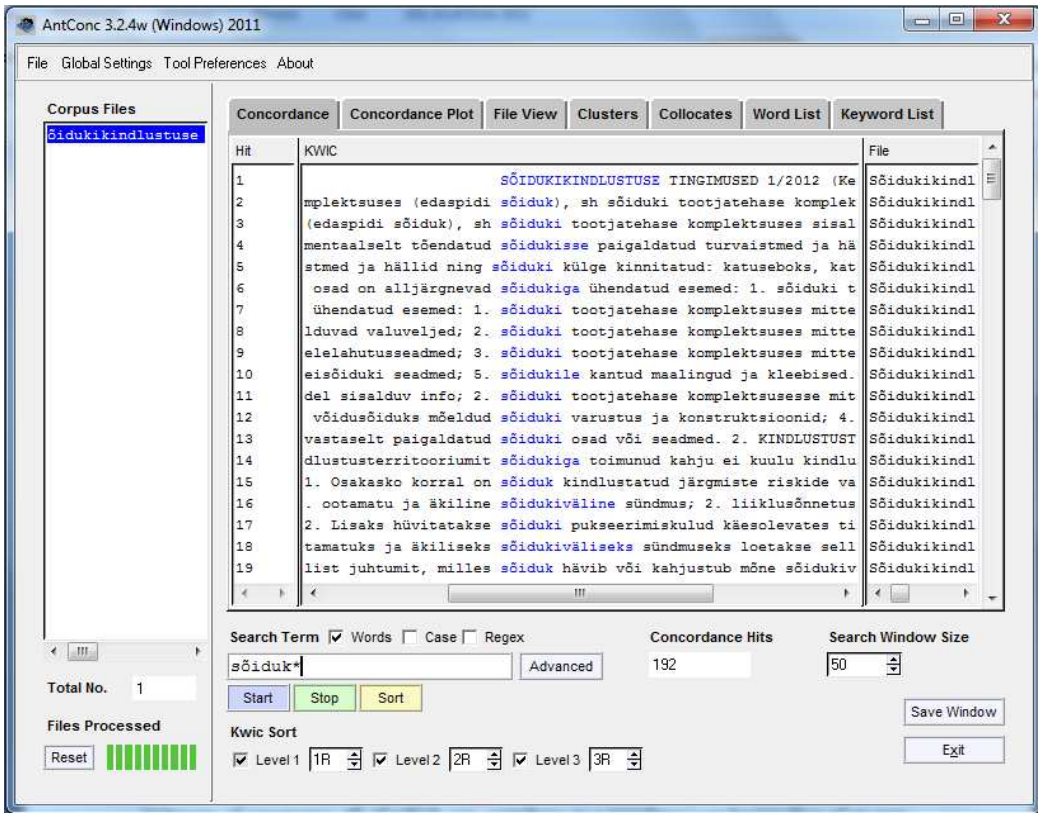


Figure 6: Retrieving different forms of the same word in the keyword-in-context list in AntConc

Clicking on **Sort** will sort the words according to what follows (see Figure 7).

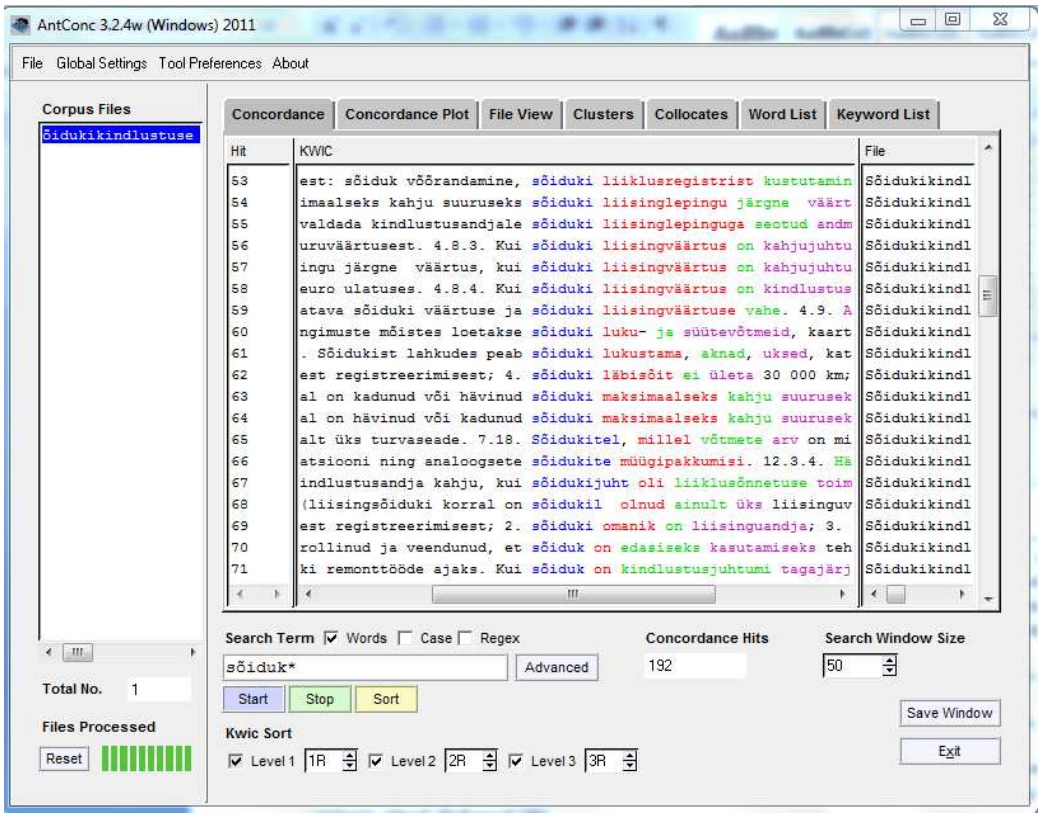


Figure 7: Alphabetically sorted concordance keyword-in-context list including all forms of the search word in AntConc

From the concordance hit list displayed in Figure 7 one can start to decide which words or strings of words should be treated as terms and need consistent translation. For that there are two options – either to apply the equivalent of the chosen term immediately to the TT in all instances where the term appears, or to draw up an ad hoc termbase. The first can be done using the Replace function, resulting in a bilingual "hybrid" text. The second can be done using, e.g, a termbase program, a spreadsheet application, or simply pen and paper. In this case, whichever method is chosen, it might be recommended to add the equivalents of the nominative of the word itself – 'sõiduk', but also its recurring collocations, such as 'sõiduki liisinguleping' and 'sõiduki liisingväärtus' (see Figure 7).

Therefore, when using AntConc, simply going through the initial list of words is not sufficient and time must be taken for searches and keyword-in-context concordance. Such work is beneficial not only because it helps to find recurring instances of terms from the ST, but also because it may reveal longer recurring multi-word units, or parts of sentences.

Continuing with the concordance, the word ‘kirjalikku’ was encountered in the word list and since it had 5 hits, the concordance list was opened. It was found that the word only appears in the phrase ‘kirjalikku taasesitamist võimaldavas vormis’. This means that the whole string can be applied at once or put in the termbase where it can later be retrieved easily.

There is also a possibility to use wildcards, receiving information only searching for a certain string of characters (see Figure 9).

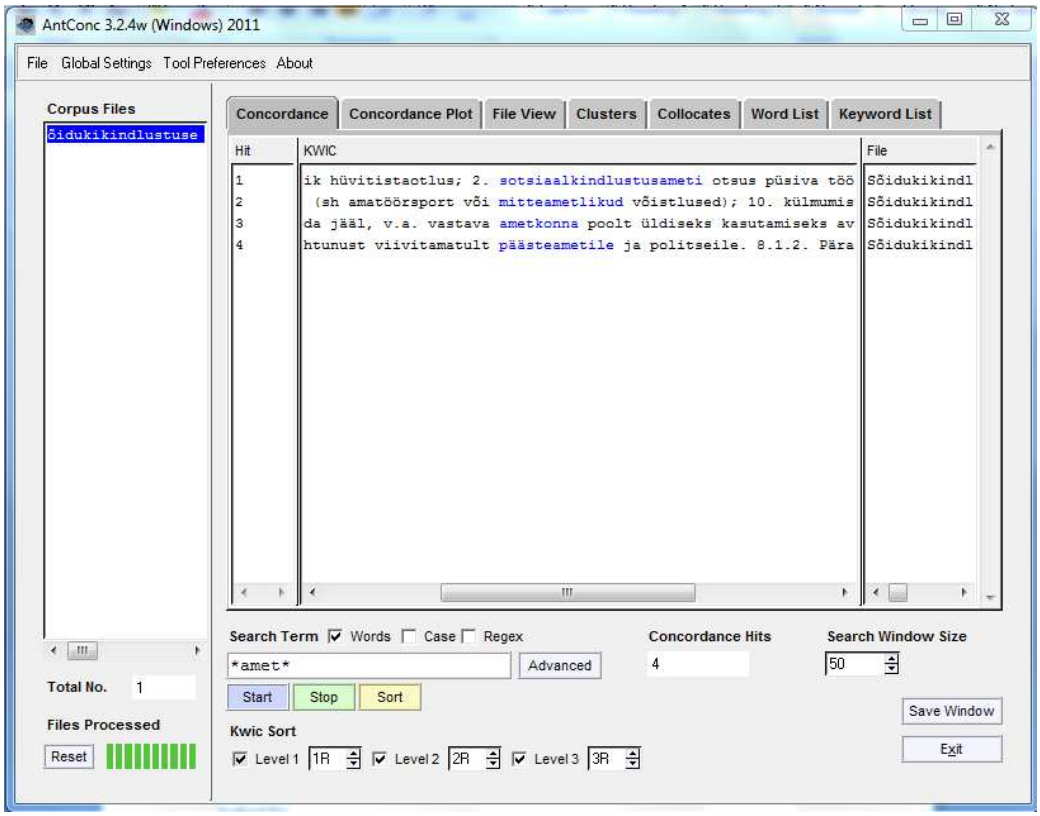


Figure 9: Searching for wildcards in AntConc

This kind of flexibility of both having an open-ended search word and getting the result as multi-word units makes AntConc a user-friendly tool that provides not only the information that the user knows to search for, but also information that the user needs but might not know to look for. Therefore, what makes a corpus-processing tool like AntConc, and especially its ability to retrieve a word list of any corpus, very useful in translating, is the fact that it will provide the user with information that they would not have known to search for without the help of the tool. The fact that one can simply read through a list of words, rather than a whole text, helps to structure the work process. Whether it also speeds up the process, can be disputed. As could be seen from Figure 1, the test translation corpus includes 1423 different words (as opposed to 4046 in the original file) that the translator must now work through. Many of these may be discarded as function words. Others need more time and patience – working with the search function and the concordancer.

Depending on the characteristics of the translator job, one may choose whether or not to pay attention to those words that appear in the text just once. In this corpus there are over 860 one-time entries. Since the corpus is in Estonian, an inflected language, a language with seemingly endless possibilities to give a word a different shape, many of these may be recurring words that we have already covered going through the first 553 entries. Not considering these one-time entries may be justified if the translation job is a one-time offer and if the translator does not expect to translate for the same client or in the same field again and does therefore not want to compile a termbase. Nevertheless, as already mentioned, the decision on whether the words occurring once in the corpus need to be studied or not, depends on the translator's expectation concerning future translations, but also the language of the corpus and whether it is inflected or not. In hindsight it may be said that the TT of the test

translation into English contains only 959 different words (compared to 1423 in the Estonian ST), i.e. approximately a third less. That makes such a concordancing tool a much easier tool to use with a less inflected language like English. However, we must also keep in mind that there are different problems arising thereof – in English verbs and nouns (e.g. dance), or nouns and adjectives (e.g. freak) can take the same form and are therefore listed under the same word in the word list. Now, if a translator would include such a word in the termbase giving it only one equivalent, and, even worse, went on to apply the equivalent of the term throughout the text, this might not simplify translating at all. However, having the opportunity of using the keyword-in-context view, the translator can decide whether these words need to receive one or two equivalents in the TT and in which instances which are used. This could not be done if the equivalent was applied to the text when meeting the term first in the process of translating. Awareness of such linguistic aspects is vital for the profession.

Admittedly, working with a concordancer before translating can be time-consuming. But so would working through a document prior translating in any other way, which may explain why most respondents said to either not read the text at all or only skim it. However, there are certain benefits to using a tool similar to AntConc – information can be received on recurrences; keyword-in-context makes it easier to find suitable equivalents; equivalents of the terms can be applied to the TT without having to pay attention to their correctness in the later editing phase. With some experience and proficiency, working through a corpus in the displayed manner will not take as long as it might seem at first glance, and learning AntConc in particular goes very fast. Eventually any user will achieve a list of terms that need consistent equivalents and are then ready to continue with carrying out the translation task without having to stop the translation flow to search for an equivalent of a term. Also, in this

manner the editing process can be more concerned with the macro-structure of the text, making sure it is TL-like and fluent.

3.2 SDL PROGRAMS

The most frequently selected CAT tool to use for translating by the respondents of the survey was a translation memory program (47%). Another 11% chose to use it together with a terminology management program. One provider of such programs is SDL, whose programs, namely SDL Trados and SDL MultiTerm, are among the best known programs on the market.

The terminology management program, SDL MultiTerm 2011, is designed for the purpose of storing and managing multilingual terminology and has a number of additional components, such as MultiTerm Extract, which makes it possible to form termbases from existing translated documents; or MulitTerm Convert, which enables to convert termbases drawn up in other tools (SDL 2011a). SDL MultiTerm, which is a separate program and must therefore be purchased separately, is recommended for use together with translation memory program SDL Trados Studio 2011. There are currently three editions of SDL Trados Studio 2011 available: Professional, Freelance and Starter Edition (SDL 2011b), the one used for this paper is SDL Trados Studio 2011 Freelance.

SDL (2011b) explains that when SDL MultiTerm termbases are accessed within the SDL Trados Studio translation environment, the source sentences are automatically searched for terms in the termbase, whereas if there are matches, these are displayed and can thereby easily be inserted into the translation. If term work has been done beforehand and a termbase is opened in the translation memory program, the active term recognition tool will pinpoint the terms in the open termbase, taking thereby the task of memorising which words need consistent translation from the user. If term work has not been done beforehand, terms must be

submitted to the termbase as they come up in translation, whereas the new entries are only useful if the translation memory program can immediately receive data thereof. An active term recognition tool is a good way to ensure consistent use of terms throughout a translated text, or even series of texts, since one termbase can be applied to many translations. The feature makes it possible to give both of the memorising tasks over to the machine.

However, most translators responding to the survey preferred to use translation memory only, without the help of terminology software. The respondents' comments lead to believe this choice is made either because the terminology program is not available to the translator or because it is found simpler and faster to work without one. Many translators also trust they remember which words they have already translated and which equivalents they chose to these. The latter can also be checked using a concordancing tool within the translation memory program. Because SDL Trados Studio is segment-based and segments, or translation units, are most likely sentences, the program does not offer the translator direct help in translating recurring terms, more so because the minimum match value recommended, and pre-set, by the developer is 70% (SDL 2011b).

3.2.1 SDL MultiTerm 2011

There are many ways for forming an ad hoc termbase in SDL MultiTerm. One of these is adding entries manually one-by-one as they come up, e.g. when working simultaneously with a concordance or while translating. But termbases can also be made by converting existing word lists (e.g. term lists in MS Excel), or by extracting terms from an existing translation.

Starting from the first mentioned – forming a termbase manually, the program should be opened from the computer program menu, and **Termbase -> Create Termbase** selected. When the **Termbase Wizard** is stepped through, the option **Add a New Entry** should be

selected from the toolbar (**Edit -> Add New** or the F3 key also fill the same function). This opens empty fields for entering the term and its equivalent (see Figure 10).

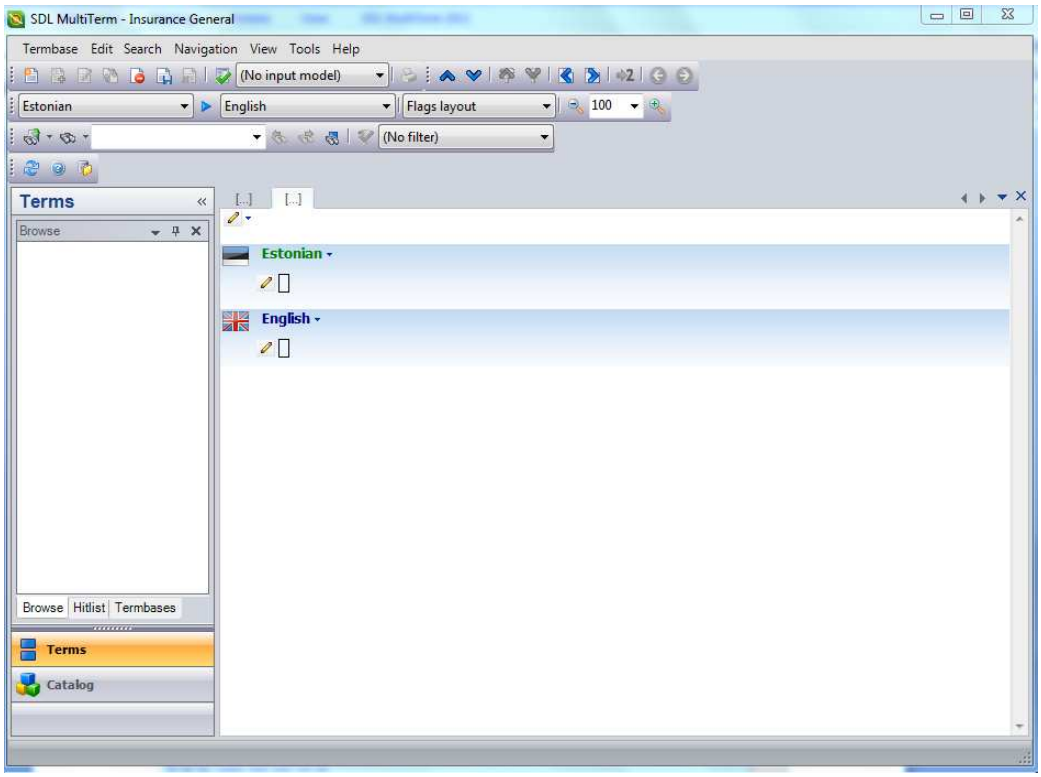


Figure 10: Empty fields for entering terms in different languages in SDL MultiTerm

Double-clicking an empty field below the language label makes it possible to type in the term. For moving to the next language field, the Tab key can be used for convenience. Before saving the entry, the Enter key must be pressed. Once those steps have been taken, the entry should be saved (by using the button **Save entry and leave editing mode** on the toolbar, the F12 key, or selecting **Edit -> Save**). The saved entries appear in an alphabetical list in the **Terms** window on the left hand side (see Figure 12). Therefore, even with the initial set-up of the termbase, the process seems fairly simple.

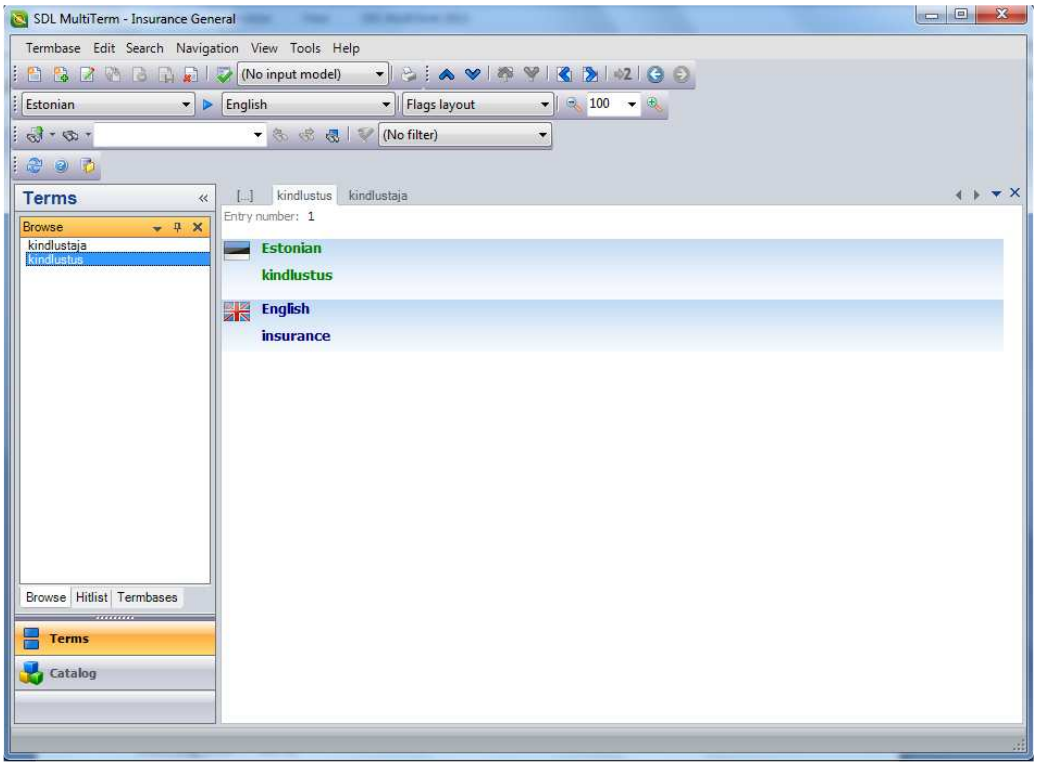
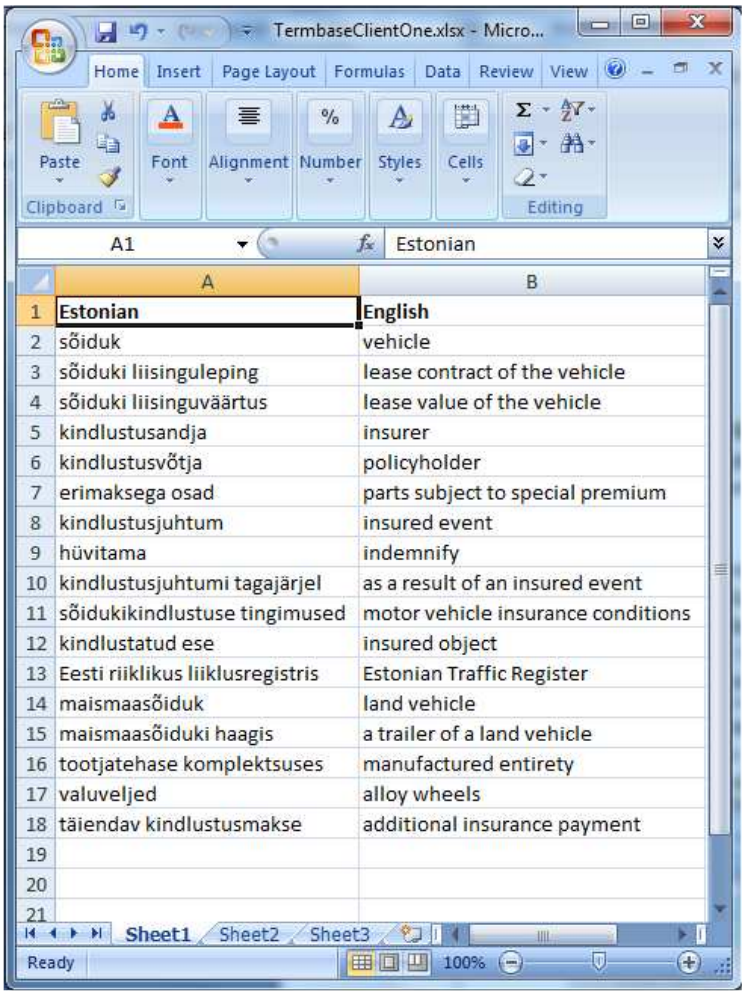


Figure 12: Saved entries in SDL MultiTerm

If an entry should be changed, e.g. for correcting spelling mistakes, after leaving the editing mode it is possible to activate the editing mode for the entry (toolbar button **Activate editing mode for the current entry**, or F2, or **Edit -> Edit**). However, what makes entering terms one-by-one inconvenient is the fact that entries cannot be saved without leaving the editing mode. This version of SDL MultiTerm has no option for adding many terms in the editing view and then saving these all together.

If a termbase has already been drawn up in MS Excel, the terms do not need to be entered into the program one by one, but can be transferred using SDL MultiTerm Converter. What is needed is a simple MS Excel file where the column headers would match the names that are suitable for fields in the SDL MultiTerm termbase (SDL 2011a) (see Figure 13).



	A	B
1	**Estonian**	**English**
2	sõiduk	vehicle
3	sõiduki liisinguleping	lease contract of the vehicle
4	sõiduki liisinguväärtus	lease value of the vehicle
5	kindlustusandja	insurer
6	kindlustusvõtja	policyholder
7	erimaksega osad	parts subject to special premium
8	kindlustusjuhtum	insured event
9	hüvitama	indemnify
10	kindlustusjuhtumi tagajärjel	as a result of an insured event
11	sõidukikindlustuse tingimused	motor vehicle insurance conditions
12	kindlustatud ese	insured object
13	Eesti riiklikus liiklusregistris	Estonian Traffic Register
14	maismaasõiduk	land vehicle
15	maismaasõiduki haagis	a trailer of a land vehicle
16	tootjatehase komplektsuses	manufactured entirety
17	valuveljed	alloy wheels
18	täiendav kindlustusmakse	additional insurance payment
19		
20		
21		

Figure 13: Termbase in MS Excel ready for conversion into SDL MultiTerm

SDL MultiTerm Convert appears as a separate program in the program menu of the computer and should be opened from there. The rest of the process, directed by a rather self-explanatory wizard, is explained in a relevant instruction manual (see, e.g. training courses provided by SDL). Once the conversion has been carried out a new termbase can be created in SDL MultiTerm, whereas its content can be imported from the converted file (under **General Settings** in the **Termbase Import Wizard**). When the wizard has been stepped through, the application switches to **Terms** view, where the terms imported from the original file are shown in a list which has automatically become alphabetical (see Figure 14). What does make

converting somewhat uncomfortable is the fact that the conversion function is not completely integrated to the program itself, but somehow added to it, and in order to actually get the termbase from the original file to the final form the user has to initiate a new program as well as step through two different wizards.

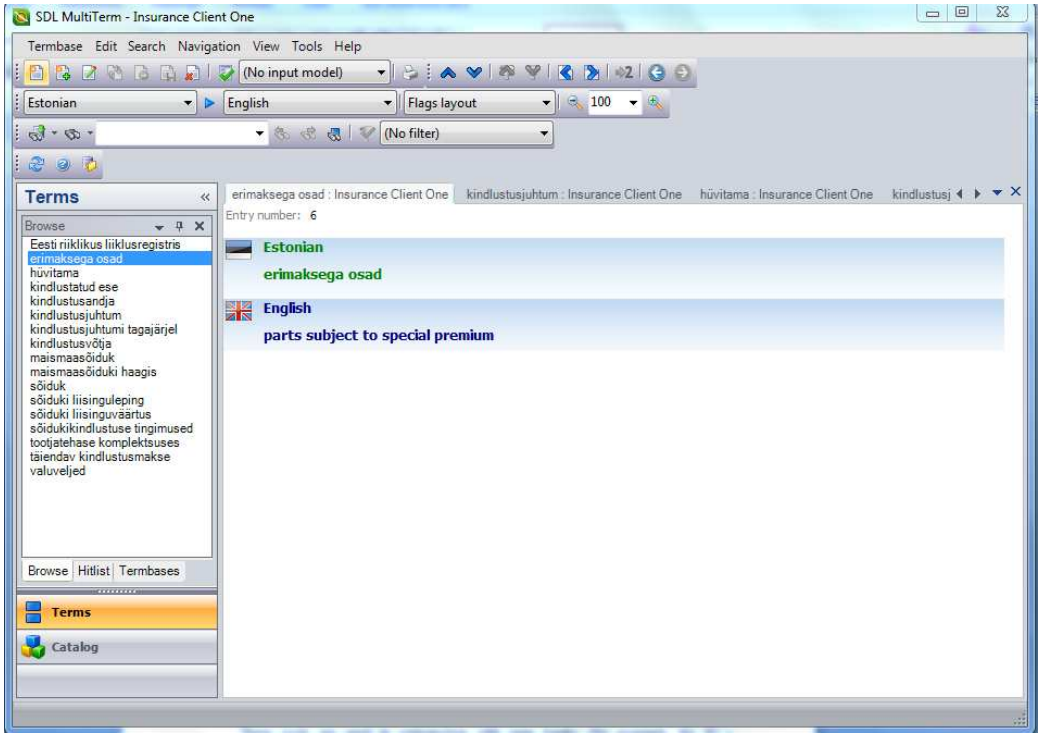


Figure 14: Terms converted from an MS Excel file to a SDL MultiTerm termbase

However, in other respects SDL MultiTerm can be considered user friendly. E.g., entries can contain more than one term for each language and in one entry various information about the term can be stored: synonyms, approved terms, obsolete terms, etc (see Figure 15) (SDL 2011a). When setting up a termbase for one particular translation task, this may seem unnecessary, but if the termbase is to be used again, additional information may prove useful. Adding information to a SDL MultiTerm termbase is much more convenient than, e.g. the MS Excel options where adding information to one term can only be done by adding a whole

column, leaving new empty fields to all the other terms. As can be seen from Figure 15, additional information stands separately for each term entry.

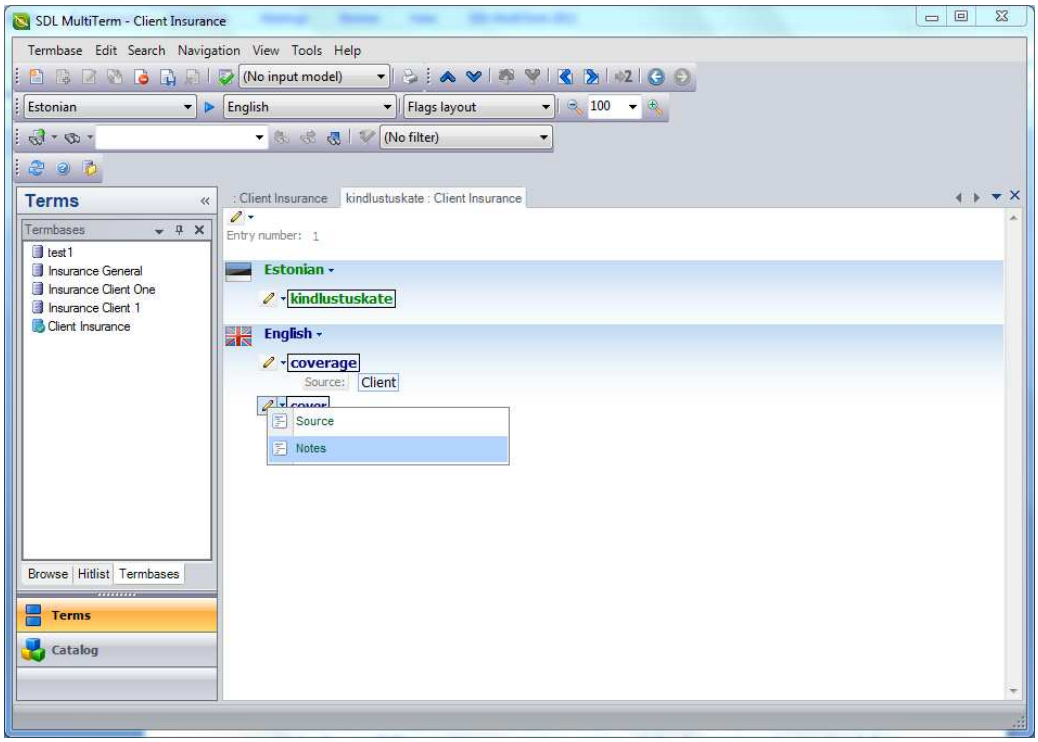


Figure 15: Adding information to an entry in a SDL MultiTerm termbase

By editing the termbase, elements (or indexes, as these are named) can be added to the termbase also after its creation. Opening the **Catalog** view on the **Navigation** pane brings up the **Catalog Categories** of opened termbases. Right-clicking on **Definition** and choosing **Edit** opens the **Termbase Wizard**, which makes it possible to edit the termbase options chosen at its initial creation. Once the wizard has been completed, the new settings are applied.

The third option for creating a termbase in SDL MultiTerm is by extracting it from a previous translation by retrieving terms as well as their equivalents from a suitable file, e.g. translation memory. Since in that case the termbase would not be created using the terms in the ST, it would not be a strict, narrow ad hoc termbase like the ones created before and would therefore be more suitable for showing how to ensure the consistent translation of terms within

the term environment rather than within one translation job. Although there is an obvious overlap, such a termbase will not be created herein. However, to explain in short how it could be done: in the case of the test translation, there would be the possibility to use the previous version of the document (i.e. conditions for 2011 and its translation) as a source for retrieving information about terms and their equivalents. Thereby, by using SDL MultiTerm Extract, a client-based rather than translation-based termbase could be formed. One must also consider that if no previous translation version of the text is available, SDL MultiTerm Extract is unhelpful since it has nothing to process. The use of SDL MultiTerm is made slightly inconvenient by the fact that some of its components, e.g. SDL MultiTerm Extract, are not integrated parts of the main program and need to be purchased and downloaded separately. To a beginner it is confusing and it seems somewhat arbitrary how some functions of the terminology management program are "inherent" to it, some are added within the initial download, and some can be added as separate applications. The highest benefits of SDL MultiTerm itself, however, can be reaped when the program used together with another – a translation memory program, where the termbases can also be modified and terms added without the need of having to open the terminology management tool separately.

3.2.2 SDL Trados Studio 2011

SDL Trados Studio, version 2011 Freelance, is a tool with many features developed for a professional translator and many study respondents also said they prefer to use Trados as their chosen translation memory program. Herein the attempt is to focus on those features of the program that are associated with translating recurring terms.

To begin with, the test translation file is opened in the program. In order to open a new file in SDL Trados Studio, information about the translation project (language direction, creation

of translation memory, etc.) is to be entered. The ST opens in the program window where it has been shown segment by segment on separate lines (see Figure 16). The segments are aligned so that they could be coupled easily with its translation – this enables an overview of the TT as well as the comparison of the TT segments with the ST segments.

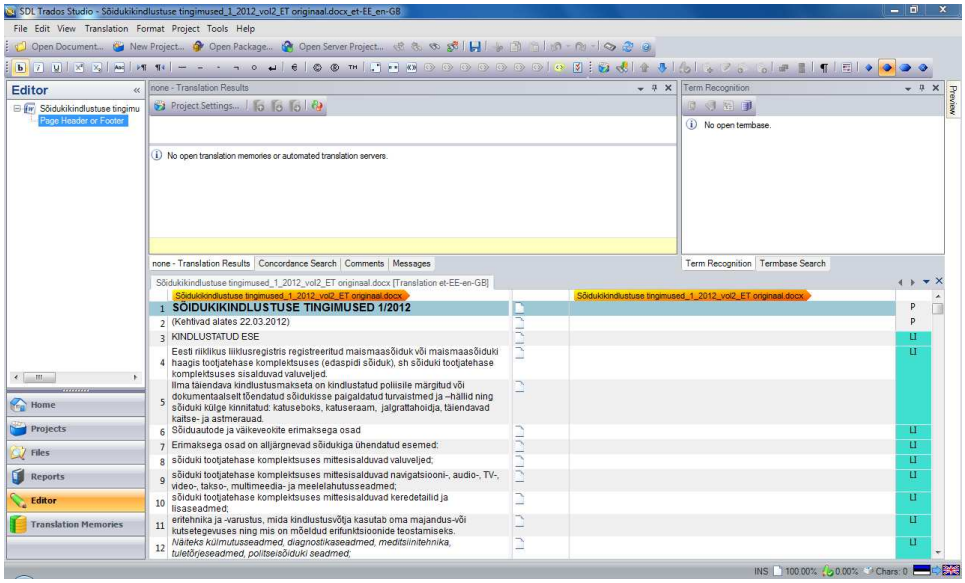


Figure 16: File to be translated opened in SDL Trados Studio (see Appendix 3 for larger view)

Moving from segment to segment is simple – by using arrows on keyboard or the mouse, the selected segment is highlighted (in blue on Figure 16). Translated sections can be confirmed by clicking on the relevant button on the toolbar (see (1) on Figure 17) or by selecting **Ctrl+Enter**. It is practical to confirm translations of segments before moving on, since that directs the segment pair to translation memory and the translations can therefore be immediately reused. That is recommended because documents themselves are often likely to contain repetitive sentences. Also, in this way the sentence is processed by the concordance feature of the program. If, however, the translator prefers to not confirm the translation of a segment, a simple file search (**Ctrl+F**) can be carried out to see which equivalents have been proposed in the working process.

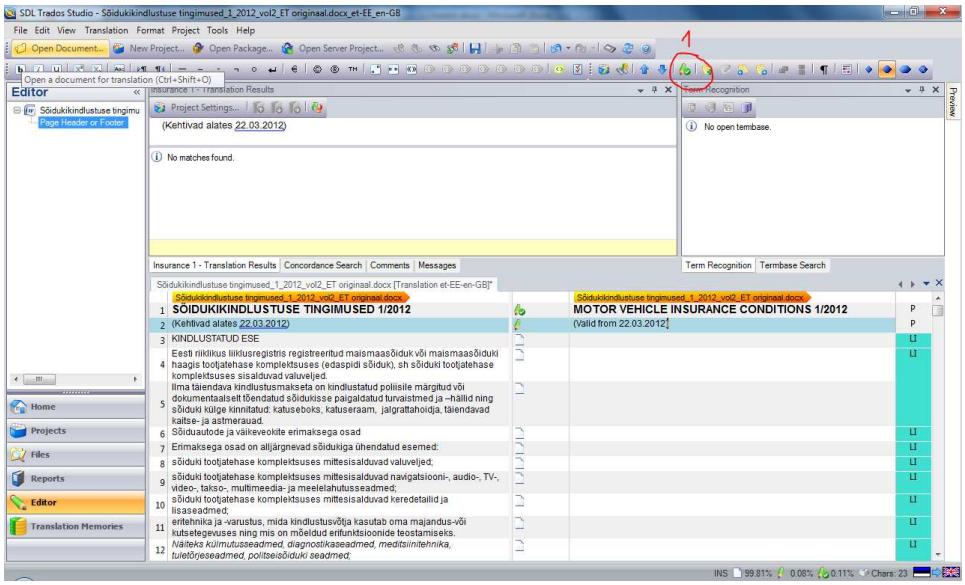


Figure 17: Confirming the translated section in SDL Trados Studio (see Appendix 3 for larger view)

Since before the test translation, no translation memory has been created, the program does not work to its full functionality. For the test translation we have the benefit that the previous version of the same insurance conditions (file to be translated) and its translation are available. The best way to create translation memory thereof is by using the alignment function, which allows the pre-processing of both the source and the target files so that their content can be imported into a translation memory (SDL 2011b). The main problem with alignment is that it will only function with full efficiency if the translated file segments correspond exactly to the source file segments – sometimes parts of text are left out or segmentation is changed during translating or later editing, in which case the alignment tool may not align the actually corresponding sentences. This can be remedied by the user before the translation memory is created, but it will recall for extra time and attention. The translator must be very careful not to create faulty translation memory pairs. To process the previously translated source and target documents SDL Trados Studio 2011 uses the SDL WinAlign component, which can be

accessed from the Home screen by clicking **Align Translated Documents** button on the toolbar (SDL 2011b).

After the alignment project has been set up, language pairs and other details as well as files have been selected, the alignment process produces an aligned file (see Figure 18).

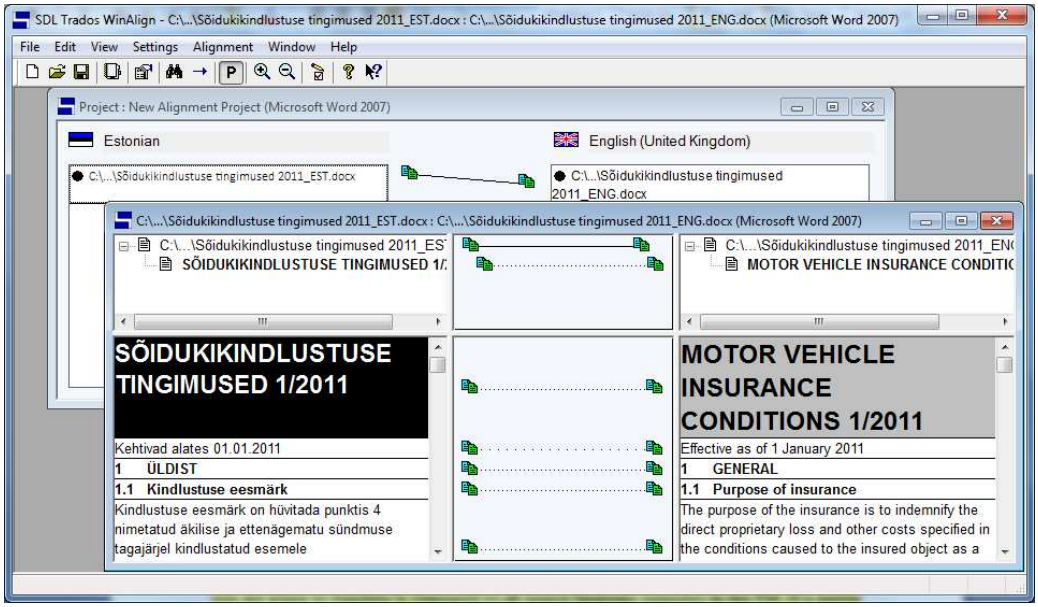


Figure 18: Alignment result in SDL Trados WinAlign, SDL Trados Studio

Next, the result must be exported to a TMX file, which can then be imported into a translation memory. Once the corresponding steps have been taken the translation memory can be seen on the **Translation Memories** screen (see Figure 19).

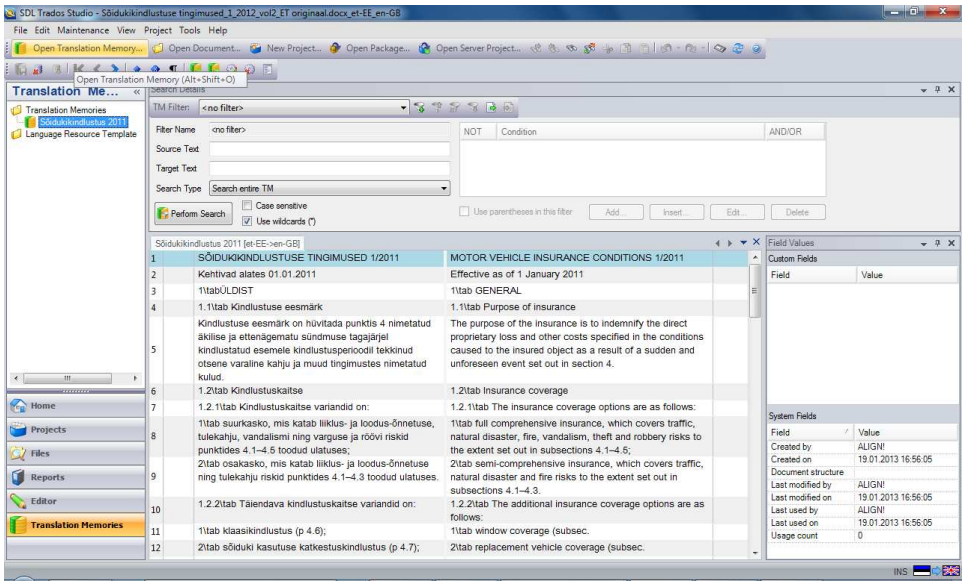


Figure 19: Aligned translation memory in SDL Trados Studio (see Appendix 3 for larger view)

Going back to the **Editor** screen, the created translation memory has become usable. Moving downwards to go through the segments, each SL segment is compared to all SL segments in the translation memory and if a similar segment, or the same segment, is found, the translation is retrieved from the memory (SDL 2011b) and displayed on the right-hand column designated for the translation (see Figure 20).

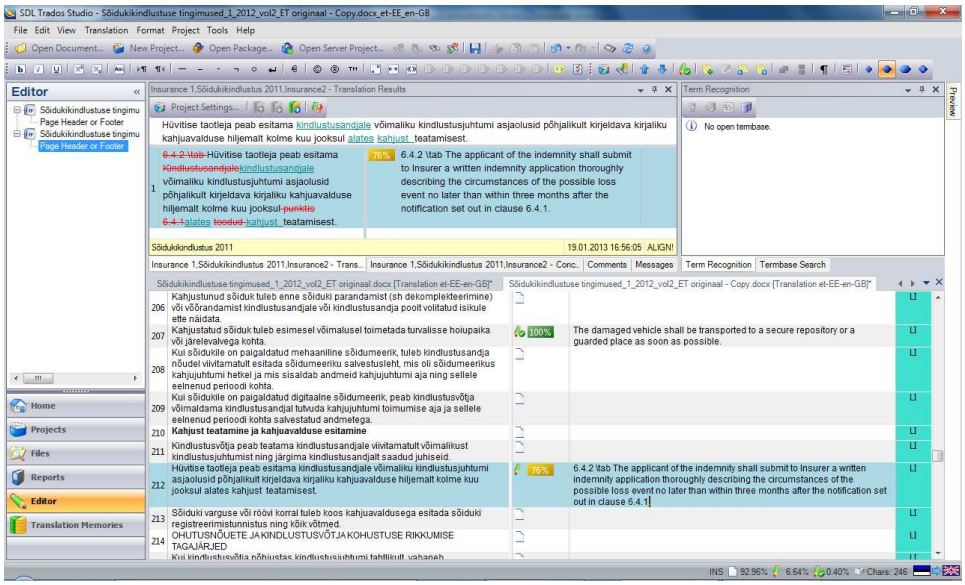


Figure 20: Translation memory results in SDL Trados Studio (see Appendix 3 for larger view)

Although SDL (2011b) states that these translations retrieved from the translation memory are displayed as suggestions, it can be seen on Figure 20 that 100% matches are automatically confirmed. Other matches, so-called fuzzy matches, are indeed suggestions. These are not completely equivalent and therefore need editing before confirming. For the ease of editing fuzzy matches, the differences between the translation memory source segment and the new translation source segment are displayed in the **Translation Results** view above the translation segments (see Figure 20). However, one must also be aware of the fact that fuzzy matches can pose challenges to the translator. First, the translator must make sure that the target language fragments that may only partially correspond to the ST segments find an acceptable translation. Secondly, the translator must also pay attention to the whole text, making sure it would not lose its coherence, a phenomenon that Bédard has referred to as the creation of a “sentence salad” – segments drawn from different sources may differ in a way that is not compatible (O’Hagan 2009: 50).

The **Translation Results** window helps the translator in making sure SL and TL segments would indeed correspond to each other. The differences are highlighted in a logical and easily comprehensible way. In matching the program does not consider variable elements such as numbers, dates or times (O'Hagan 2009: 48); if a match is found that differs by only a variable, the translation suggestion appears in the TT section with the variable from the new ST and not the translation memory ST.

The test translation revealed also that there is a possibility that a segment actually appears in the translation memory, but does not show as a suggestion. E.g. the title 'kindlustatud ese' does not appear as a suggestion. Since not receiving a memory suggestion for the title raised suspicion, a concordance search was carried out. Menu item **Translation** -> **Concordance Search**, or F3, opens the relevant view above the translation segments, where a search term can be put in (see Figure 21).

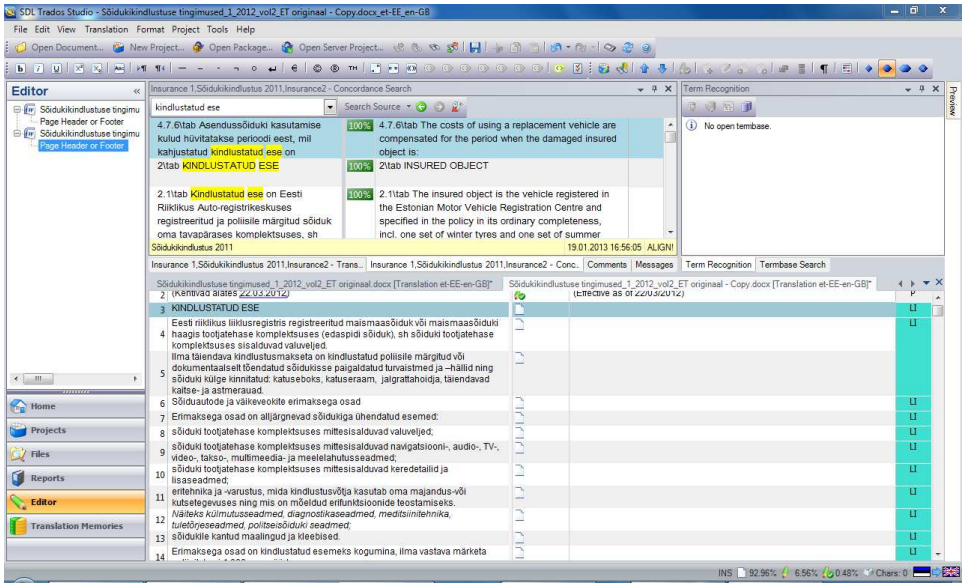


Figure 21: Concordance search in translation memory files in SDL Trados Studio (see Appendix 3 for larger view)

This search reveals that there indeed is a similar source segment in the translation memory. The reason why it did not appear in the translation view as a suggestion may be due to the fact that there are some formatting-related differences (for some reason the aligned translation files show tabs as text items) and these form too high a percentage of the translation segment. As Bowker (2002: 200) explains, fuzzy matching mechanisms use character-based similarity metrics where resemblance of all characters in a segment is checked, whereas all characters include also punctuation, etc.

This suggests that it might be advisable to lower the minimum match value, which can be done in **Project Settings**, under **Translation Memory and Automated Translation -> Search** (see (2) on Figure 23 on page 74). Once the minimum match value has been changed to 50%, a translation suggestion for ‘kindlustatud ese’ is offered (see Figure 22).

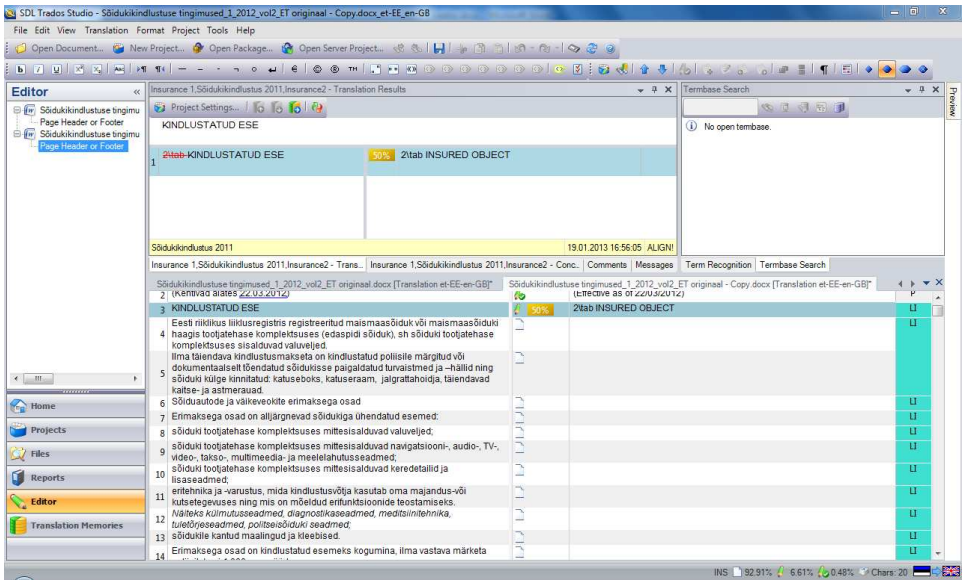


Figure 22: Translation suggestion appears after lowering minimum match value on SDL Trados Studio (see Appendix 3 for larger view)

SDL has preset a 70% matching. Some have suggested that matches as low as 40% can be useful (O’Brien 1998: 117). When thinking of retrieving information on how certain terms have been translated before, this may indeed be so since there is a higher chance of being

offered a previous translation. Kenny (2011: 468) argues that there is no need to lower matching percentage since only very high-value matches actually save time for translators. By that Kenny takes the more pragmatic approach whereby meeting a deadline is valued over achieving a better result. That is, assuming, that a lower matching percentage helps the translator find term equivalents and results in a more consistent TT. The lower the matching percentage, the higher the number of previously translated words or stings of words would be brought up and offered for reuse. If the translator decides to take the time to see these offered low matches through to find parts that could be reused, more recurring terms could be identified. However, Bowker (2005: 18) has also concluded that translation memory does not always improve consistency and for translators it is hard to balance between productivity and quality, since often one comes at the expense of the other. With regard to term consistency, it is important for the translator to pay attention to which source the offered match is found from – e.g., if it is the same text that is being translated, no further analysis may be necessary; but if it is a text translated by some other translator some time ago, analysis on the suitability to reuse the equivalent should be more thorough. The viewed program, SDL Trados Studio, does not enable to do this by default, but there is a possibility to initially set the software so that an additional field with the document name where the information is retrieved from would be added to each translated segment (Kļaviņš, personal communication 10 May 2013).

Although using translation memory is advertised to facilitate achieving uniformity in translation in a way, using translation memory only, even together with the concordance search, does not necessarily ensure the consistency in term translation. On the contrary – it may be the case that the translation suggestions retrieved from memory include different equivalents and therefore different equivalents will be entered into the TT. The main issue

with the concordancing feature is the fact that it presumes the translator remembers a certain term has already occurred in the translated text and therefore knows to check its previous translation.

One possibility for avoiding such inconsistencies is using a translation memory tool together with a terminology management tool. To avoid a situation whereby translation memory segment translations and terminology tool entries bring about different equivalents to terms, the translation memory concordancer can already be used when deciding upon an equivalent for a term in a termbase. If in the termbase no equivalent has been entered for the term ‘kindlustatud ese’, the concordance result can be used for entering the translation memory equivalent into the termbase as well as into the translation. For that the termbase should be opened in SDL Trados Studio. That can be done in **Project Settings**, where under **All Language Pairs** we find the **Termbase** options (see (1) on Figure 23).

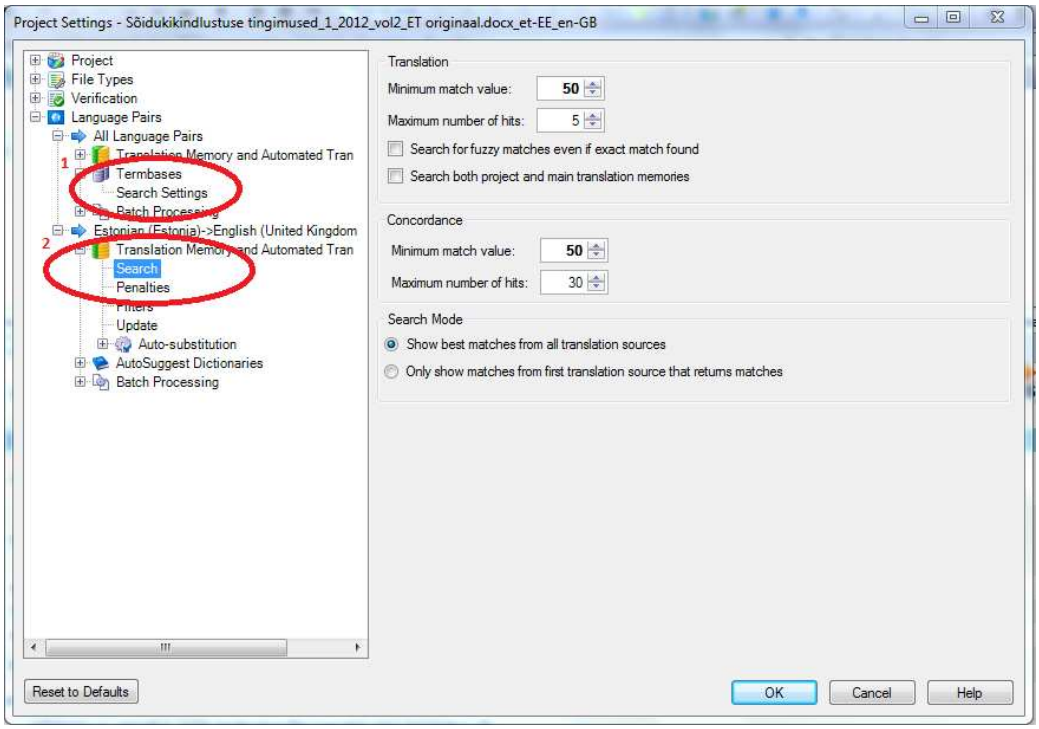


Figure 23: Adding a termbase; and changing minimum match value in Project Settings in SDL Trados Studio (the latter option activated)

Once all suitable termbases have been selected, these become available in the **Termbase Search** window on the right-hand side above the translation segments. When a search word is entered, the results are displayed in the same window with all matching terms (see (1) on Figure 24).

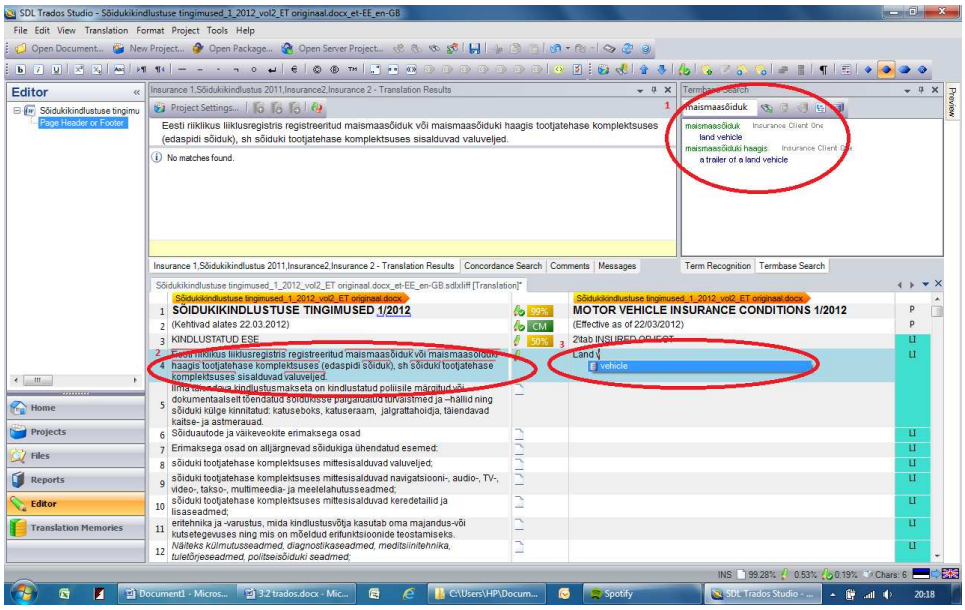


Figure 24: Using a termbase in SDL Trados Studio (see Appendix 3 for larger view)

In the marked ST segment all terms that have been entered into any of the opened termbases have been highlighted with a red bracketed line above the word or phrase (see (2) on Figure 24). This suggests that the Active Terminology Recognition function is operating. It alerts the translator to pay attention to those words since these are the ones the translator has decided to add to the termbase. Like this the translator is no longer dependent on personal memory and no longer will mistakes from forgetting that a certain word is indeed to be treated as a term going to cause inconsistency.

As the translator continues with the translation, target term suggestion application is switched on and as the first letter of a term appearing in the segment is typed in, its equivalent is offered from the termbase (see (3) on Figure 24). Instead of typing the word, the translator can now simply press the Enter key and the suggestion is inserted (SDL 2011b). All SDL MultiTerm entries can be modified or viewed in full in the same **Termbase Search** view by simply right-clicking on the term under interest. However, when editing a termbase entry by changing the equivalent term, one must be aware of whether the term has already been used in

the text or not. Also, it is worth noting that the active terminology function points out fuzzy matches as well, e.g. if the term has been defined in the termbase as ‘täiendav kindlustusmakse’, but the text includes words ‘täiendava kindlustusmakseta’, the latter is still highlighted (view segment 5 in Figure 25). The likelihood that a term is overlooked due to grammar, or even spelling mistakes in the TT, is thereby minimised.

SDL Trados Studio has another feature that improves not only the speed of translation, but also, ideally, the consistency of the translated text. That feature is an AutoSuggest dictionary, which, if opened with a translation project, matches segment fragments to dictionary entries and offers these entries to the translator for acceptance while the translator is typing. If the dictionary suggestions are indeed correct, this feature does doubtlessly improve translation productivity. However, SDL (2011b) themselves accept that these dictionary matches are usually not as reliable as terminology suggestions from termbases, as these are created from translation memory and not maintained by the translator. In order to use AutoSuggest dictionaries, these must first be created from existing translation memories. However, the problem is that SDL has considered this to be an extra feature of SDL Trados Studio and has included it only with the Professional or Freelance Plus editions of the software. Since the test translation is carried out with the Freelance Edition, the effectiveness of this feature cannot be analysed further herein.

After having just praised the active terminology recognition function in ensuring the consistent translation of terms and knowing that most translators choose to use a translation memory program without a terminology management tool, the question arises – how to best use SDL Trados Studio in ensuring term consistency without SDL MultiTerm? If there is no possibility for automatic term recognition, term replacement could be tried instead. For that,

the ST can be copied to the TT fields using the **Copy All Source to Target** function (in **Translation** menu, or Atl+Shift+Ins) creating thereby a version of the ST that can be modified (see Figure 25). Note that already translated or edited segments were not changed, so if these have not been previously checked against an ad hoc termbase, inconsistencies may remain therein.

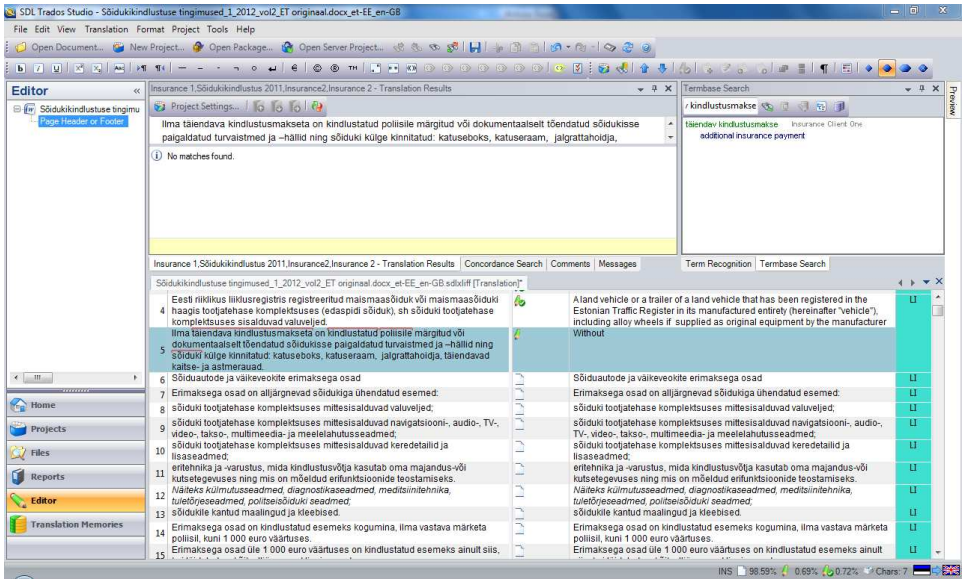


Figure 25: ST copied into TT fields in SDL Trados Studio (see Appendix 3 for larger view)

Now terms could be picked from the termbase or as they come up in the text and applied throughout the TT. That can be done by opening the **Find and Replace** window and entering the term and its equivalent (see Figure 26). Note, that in inflected languages, such as Estonian, it is better to use the stem of the word or search for various forms. When wishing to separate the word ending from the inserted term, any symbol could be entered.

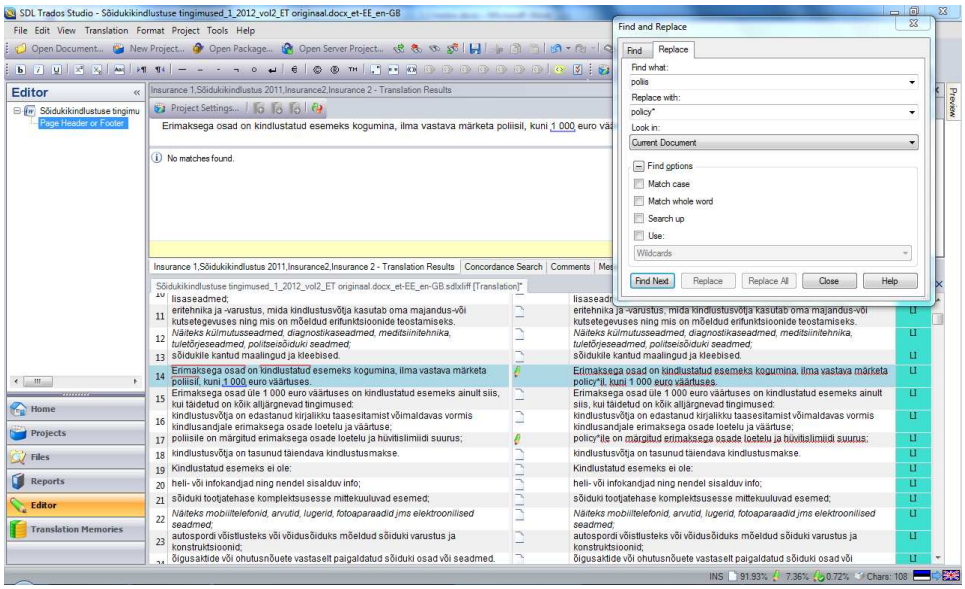


*Figure 26: Applying terms using **Replace** function in SDL Trados Studio (see Appendix 3 for larger view)*

It is worth noting that segments where changes were made have now been marked as edited. Once all terms have been entered to the TT in such a manner the text will become bilingual. However, thanks to the layout of the program the original ST segment is right next to the new bilingual version which can now be used as a kind of a termbase of its own, provided exclusively for the segment in question. Should a term come up later during translating and there is doubt whether it has appeared already before in the text, there are, again, two options for checking. One of these is to insert that term in the **Concordance Search** (F3), which will also show the occurrences of the word in previously translated files (i.e. translation memory). However, it will not show any unconfirmed translations in the translation file. Therefore the **Find** function could be used. Note that the cursor should be on TT field. When looking for the term ‘erimakse’ from the ST, **Source** should be selected in the **Find** window. Also, it might be a good idea to make it a wildcard search, since that displays also options we might not have considered. Therefore, ‘makse’ (see (1) on Figure 27) was looked for in the ST (see (2) on Figure 27) as a wildcard (see (3) on Figure 27). The first

match is then marked in red (see (4) on Figure 27) even if it only forms a part of the total word. Once it has been marked somehow in the TT, the translator can move on by clicking **Find Next** in the **Find** window.

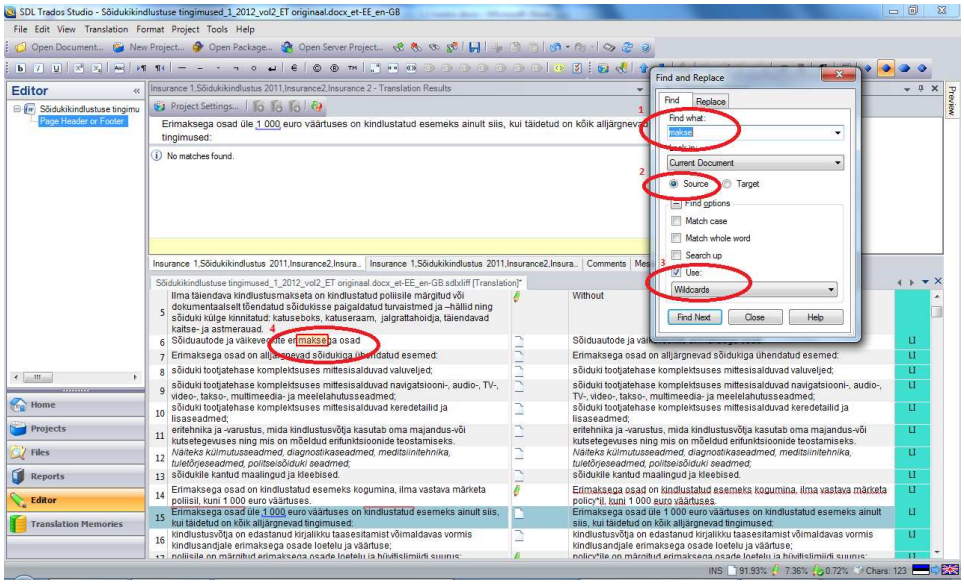


Figure 27: Finding terms in the ST in SDL Trados Studio (see Appendix 3 for larger view)

Better term consistency is not provided by the program without the translator's attentiveness and know-how on what to pay attention to. But training programs are costly and the learning curve "steep" (O'Hagan 2009: 50). SDL provides user training on different skill levels and has produced thick and detailed training manuals explaining the many and varied features of the program. However, and this may explain why some respondents considered Trados and similar programs to be easy, one can start to reap the benefits of a translation memory program with only knowing a handful of functions.

3.3 MICROSOFT WORD

Altogether 40% questionnaire respondents answered to choose MS Word or a similar word processor for translating the text described in the translation case. That means that a considerable amount of translators do not use the help of translation memory tools in their

work. However, they must also be able to provide the client with a satisfying result. Can it also be guaranteed that recurring terms receive only one equivalent if the text is translated in MS Word?

When starting to translate in MS Word, the translator has many options how to approach the work – whether to open a new file and type the translation in there (with or without formatting), whether to write the translation over the original text, whether to align the translation next to the original (paragraph by paragraph or sentence by sentence, etc.), etc. The study made a difference between two options: (i) new file, and (ii) on top of or next to the ST; because the two can be differentiated by ability to apply terms throughout the text. Most respondents who use a word processor chose the latter option. However, it is likely that in making their decision they had in mind the convenience of applying formatting or the general ease of use (comparability of ST and TT) or other issues, rather than the consistency of terms.

When coming across terms in the text, the translator again has options – whether to translate the term only in one place or apply the equivalent also elsewhere in the text. The study showed that translators do not tend to apply the equivalents of the terms throughout the whole text. This option would, of course, not be available to those who prefer to open a new, empty file when starting to translate, and that is why the option can be considered the least flexible and least helpful in working with recurring terms. However, when the translation is written into a file that includes the whole original text, the TL equivalents of the terms can be entered into the ST so that when translating the text, the translator can already see the TL word in the ST. This results in a “hybrid text”, which is more difficult to treat the more terms there are that have been replaced. However, if such replacing is made keeping in mind that lexical items should not be replaced, but only content-words, the text should be better readable. After

making some replacements of words collected previously to a MS Excel termbase with the help of AntConc, the text did look rather confusing, following the logic of one language with the words of the other (see Figure 28).

1. INSURED OBJECT
   1.1. Eesti riiklikus liiklusregistris registreeritud maismaavehicle- või maismaavehicle-i haagis in manufactured entirety (edaspidi vehicle-), sh vehicle-i in manufactured entirety sisalduvad alloy wheels. Ilma täiendava kindlustusmakseta on kindlustatud poliisile märgitud või dokumentaalselt tõendatud vehicle-isse paigaldatud turvaistmed ja –hällid ning vehicle-i külge kinnitatud: katuseboks, katuseraam, jalgrattahoidja, täiendavad kaitse- ja astmerauad.

Figure 28: Text after replacements

In such a case a split-screen option can be used in order to keep the “pure” original next to the “modified” original version to ease translation. However, in that case, the translator must go between two different locations to get all necessary information for translating a sentence and may therefore not be pleased with the method. There is another option – replacements can be made before or after the SL term, keeping the SL term in the text and providing its equivalent (see Figure 29).

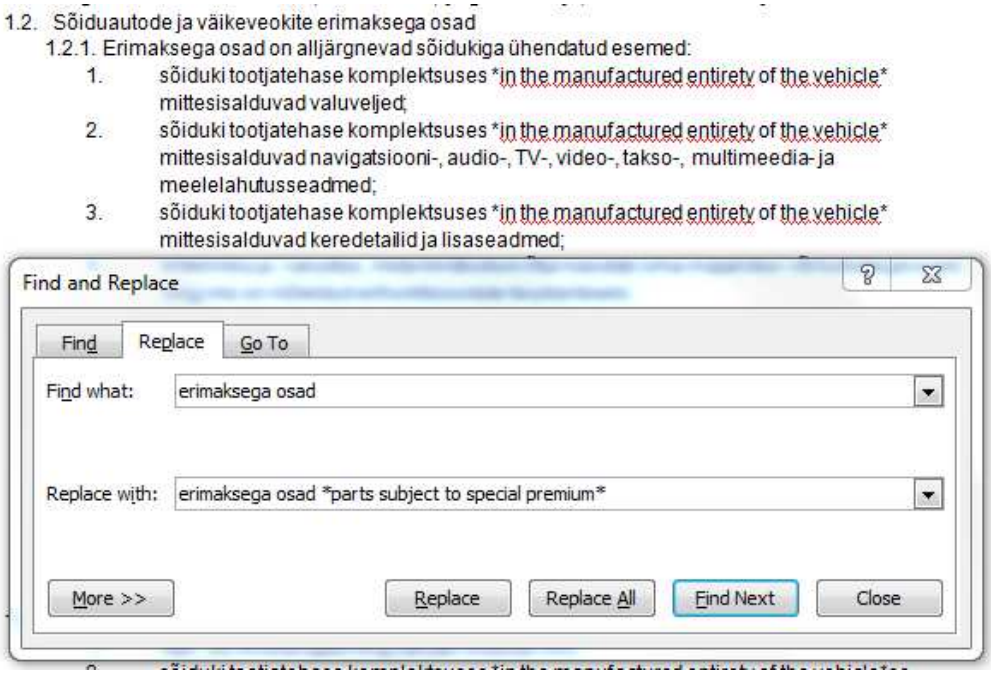


Figure 29: Adding equivalent with keeping the original

In this way the translator will be automatically informed of the fact that the word or string of words is a term and of what its equivalent is; i.e. both memorising tasks are given over to the computer. The text will become longer and is still harder to read than a one-language text, but since no parts have been “cut off”, there is no actual need for comparing the hybrid text to the ST in a separate or split-screen window.

Applying terms in the described manner can be done either (i) before starting to translate, or (ii) during translation, in which case the translator must take care that all terms are immediately applied. The later version stops the translation flow more frequently and may therefore be considered less convenient. However, the first option assumes working with terms prior to translating. Only a marginal number of translators admitted to prefer working in such a manner.

Some respondents said they prefer editing to pre-work when it comes to recurring terms. Since that kind of editing must be ST-based, there are two ways how this can be done. One option is to create a split screen where the search window is opened in the ST file and then the TT file is scrolled to compare all instances of the word found in the first window. The other option is to keep both the ST and the TT in the translation file until that stage and go through terms with the search function. In the latter case the search may be somewhat faster since there is only one file to operate with. In both cases, the search must be carried out in the SL first and then all equivalents must be seen through in the TT, since looking for the TT equivalent such instances of the terms where typing mistakes or synonyms appear, will not be found. Either way, one must make sure the "Match Case" option would not be used to find all instances of the word. Also, the "Use wildcards", "Ignore white-space characters", and "Find all word forms" could also be helpful in finding different forms of the same word or words with spelling or typing mistakes (see Figure 30).

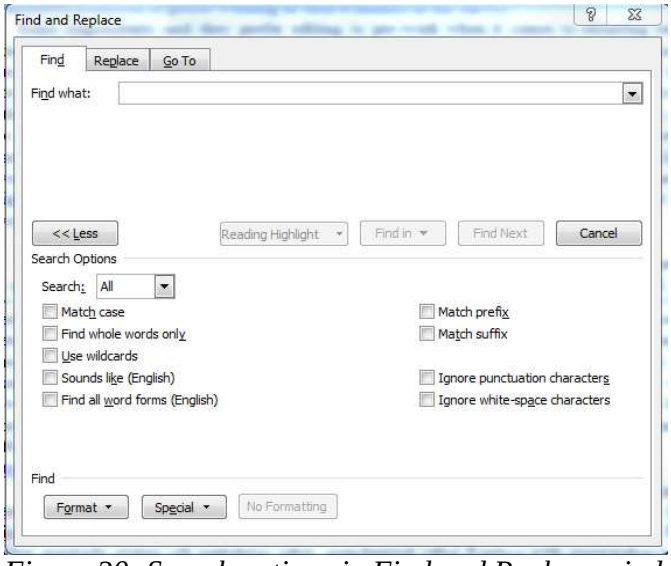


Figure 30: Search options in Find and Replace window

However, one must also acknowledge that if the ST is in many files and the translator works linearly, translating one file first and second file later, terms may get different

equivalents if a termbase is not drawn up and if the translator does not remember to check all the terms from that termbase. Opening both files at the same time and applying all terms into both files with the Replace function would enable to keep from that, but may be considered somewhat more inconvenient. Same applies for later editing. To ensure all terms in both files get the same equivalent, four files would have to be open at the same time – both ST and both TT files, whereas the search would have to be made in two files that are compared then to the other two. If a word list had been withdrawn from both files before translating (like it can be done in AntConc), and all equivalents applied immediately to both files, editing recurring terms would not be necessary. However, as could be seen from before, this order of tasks is also somewhat time-consuming and the final selection of which method of work is most convenient is for each translator to make.

3.4 CONCLUSION TO SOFTWARE DEMONSTRATIONS

In this paper, no one speedy and simple method for achieving precise and consistent results that would have been proven to be better than other methods in giving recurring terms only one equivalent throughout a translation job can be proposed. However work methods have been displayed in this chapter that could help translators to structure their work with special attention to the issue of recurring terms depending on which program they choose use.

The program viewed first, a concordancer, enables ST analysis that would give the translator information about recurring terms and other text elements based on which the translator can form an ad hoc termbase or apply the equivalents of recurring terms in the text immediately. The viewed termbase program, SDL MultiTerm, is efficient when used together with a translation memory program, SDL Trados Studio, as then it can also facilitate the

translator's work by taking up an additional memorising task. But when the termbase program is used alone, it is able to take up only one of the memorising tasks and is therefore equal to other kinds of termbases, e.g. in a spreadsheet application or on paper. A translation memory program is helpful when used together with a terminology program, but can successfully be employed also on its own, as even in this case the translator can be aided by the program in both memorising tasks. Also, it must not be forgotten that such a program enables the additional benefit of ensuring term consistency within the larger text environment. But where consistency is to be ensured within one document, a word processor is equivalent to a translation memory program, even if the used functions differ somewhat.

All in all, the software demonstrations showed that it is not so much the tool we use that makes a difference, but how we use it – meaning that ultimately it is the translator who takes responsibility for the result of their work and that a good and attentive translator, who knows how to best employ their tools, should be able to achieve good results independent of which CAT tool they use.

In this chapter specific options were offered to translators in regard to how to make their work with recurring terms more technical and systematic and perhaps therefore also somewhat more convenient. Hopefully translators welcome this kind advice and want to test some of the offered methods. However, herein the author can only encourage translators to try new ways of work, without claiming the ones included in this paper to bring any better results than the methods already used.

CONCLUSION

Making sure a recurring term receives the same equivalent throughout a translated text is only one aspect of many that a translator has to keep in mind when working on a commissioned text. The importance of it derives from many factors, avoiding confusion in reading the document and achieving the client's satisfaction being among these. As discussed in the paper, clients rate consistency in terminology highly – they consider a bad translation one where the use of terms is inconsistent, where their provided dictionaries or the terminology usage established in the target language has been ignored. As service providers, translators should keep this in mind.

This thesis has attempted to show how translators can adhere to such demands without having to change their work methods to a considerable extent. First a survey was conducted to find out translators' preferred work methods in as much as these are connected to the translation of recurring terms. As it appeared, all respondents use some kind of a CAT tool in their work. Based on that knowledge, software demonstrations were carried out with a view of displaying which methods can be used within the chosen programs to facilitate the translation of recurring terms. In addition to more common tools, a concordancer, a program that is not widely known to be used for source text analysis, was also introduced as a method for structuring work with recurring terms. The tool facilitates receiving information on recurrences in the source text. The aim of these demonstrations was to offer translators ideas on how to structure their work differently or what to change in their work practices to be more knowledgeable in ensuring term consistency. Encouragingly, the study results also showed

that translators, especially those that have studied computer programs before, welcome a chance to learn more about the tools they already use.

However, it is important to understand that all of the covered CAT tools are only used to aid the human translation and the responsibility for the result will be solely the translator's. Therefore, translator awareness of the importance of the covered subject as well as knowledge of the possible methods to use in their work is paramount.

The next step in research would be to find out whether the interrelationship between the translator and technology is functional in a way that more advanced use of CAT tools by which both memorising tasks are given to the computer, enable translators to achieve better results concerning term translation consistency or not. Other factors that may influence the result are translator attitude and commitment; how meticulous the translator is in their work and how willing to spend time on this particular aspect of translation. In this study many survey respondents were found to prefer choosing speed and simplicity of work methods over such an aspect of translation as the consistent translation of recurring terms. It might be worth investigating whether they have made the right decision.

Judging simply by how easy it is to find inconsistencies in term translation in translated texts, the issue should be concerning. This paper hopes to reach translators and raise their awareness on the matter, but also to inspire translators to review their work methods with recurring term consistency in mind, to make them think if they really do make sure terms only get one equivalent in the target text, or if they should change something in their work methods to ensure that.

REFERENCES

Aston, Guy. 2009. Foreword in Beeby, Allision and Patricia Rodriguez Ines, Pilar Sánchez-Gijón (eds.) *Benjamins Translation Library. Corpus Use and Translating*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Baer, Brian James and Tatyana Bystova-McIntyre. 2009. Assessing cohesion: developing assessment tools on the basis of comparable corpora in Angelelli, Claudia V. and Holly E. Jacobson (eds.) *Testing and Assessment in Translation and Interpreting Studies.* Amsterdam and Philadelphia: John Benjamins Publishing Company.

Bowker, Lynne. 2002. Computer Aided Translation Technology: A Practical Introduction. Ottawa: University of Ottawa Press.

Bowker, Lynne. 2005. Productivity vs quality? A pilot study on the impact of translation memory systems. *Localisation Focus 4*, March 2005. Available for purchase at http://www.localisation.ie/resources/locfocus/pdf.htm, accessed 29 April 2013.

Dictionary.com, LLC. 2012. Available at http://dictionary.reference.com/browse/irregular?s=t, accessed 20 December 2013.

Goldblatt, Howard. 2002, April 28. *The Writing Life.* Washington Post. Page BW10. Available at http://old.dongxi.net/b07DG/print, accessed 21 December 2013.

Hatim, Basil and Jeremy Munday. 2004. *Translation. An advanced resource book.* London and New York: Routledge.

Kaska, Veronika. 2011a. *Eesti viisapoliitika rändekanalina*. Tallinn: Sisekaitseakadeemia. Available at http://riksweb.sisekaitse.ee/index.asp?action=102&tid=51979, accessed 19 January 2013.

Kaska, Veronika. 2011b. *Visa policy as migration channel in Estonia*. Tallinn: Sisekaitseakadeemia. Available at http://riksweb.sisekaitse.ee/index.asp?action=102&tid=51979, accessed 19 January 2013.

Kenny, Dorothy. 2011. Electronic Tools and Resources for Translators in Kristen Malmkjæer and Kevin Windle (eds.) *The Oxford Handbook of Translation Studies*. London: Oxford University Press.

Koby, Geoffrey S. and Brian James Baer. 2003. Task-based instruction and the new technology: Training translators for the modern language industry. In Brian James Baer and Geoffrey S. Koby (eds.), *Beyond the Ivory Tower: Rethinking Translation Pedagogy*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Legaltext. 2013. *Law of Obligations Act.* Available at http://www.legaltext.ee/et/andmebaas/paraframe.asp?loc=text&lk=et&sk=en&dok=X3008

5K6.htm&query=v%F5la%F5igusseadus&tyyp=X&ptyyp=RT&fr=no&pg=1, accessed 25 January 2013.

Malmkjæer, Kristen. 2005. *Lingiustics and the Language of Translation*. Edinburgh: Edinburgh University Press.

O'Brien, Sharon. 1998. Practical experience of computer-aided translation tools in the software locatlisation industry. In Bowker, Lynn, M. Cronin, D. Kenny and J. Pearson (eds.). *Unity in Diversity? Current Issues in Translation Studies.* Manchester: St Jerome.

O'Hagan, Minako. 2009. Computer-aided translation (CAT) in Mona Baker and Gabriela Saldanha (eds.) *Routledge Encyclopedia of Translation Studies*, Second Edition. London and New York: Routledge.

O'Hagan, Minako. 2013. The impact of new technologies on translation studies: A technological turn? In Millan, Carmen and Francesca Bartrina (eds.). *The Routledge Handbook of Translation Studies.* London and New York: Routledge.

ÕS = Eesti õigekeelsussõnaraamat. 2006. Tallinn: Eesti keele sihtasutus. Available at http://www.eki.ee/dict/qs/, accessed 15 January 2013.

Pym, Anthony. 1993. Available at http://www.erudit.org/revue/ttr/1993/v6/n2/037160ar.pdf, accessed 15 January 2013.

Rodríguez Inés, Patricia. 2009. Evaluating the process and not just the product when using corpora in translation educaiton in Beeby, Allision and Patricia Rodriguez Ines, Pilar Sánchez-Gijón (eds.) *Benjamins Translation Library. Corpus Use and Translating*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Saluäär, Hille. 2007. Mis on tõlkepõhine tekstianalüüs ja mis kasu on sellest eesti õigustõlkijal? *Õiguskeel* 2007/4. Available at http://www.just.ee/33353, accessed 13 January 2013.

Sánchez-Gijón, Pilar. 2009. Developing documentation skills to build do-it-yourself corpora in the specialised translation course in Beeby, Allision and Patricia Rodriguez Ines, Pilar Sánchez-Gijón (eds.) *Benjamins Translation Library. Corpus Use and Translating*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

SDL Plc. 2011a. Training Course: SDL MultiTerm 2011 for translators and project managers. Available for purchase on www.sdl.com, accessed 10 January 2013.

SDL Plc. 2011b. Training Course: SDL Trados Studio 2011 Getting Started for Translators. Available for purchase on www.sdl.com, accessed 10 January 2013.

Sutcliffe Moraes, Naomi J. 2008. The Translator's Binoculars, Part 1: How to Search Monolingual Reference Texts in *The ATA Chronicle*, Volume XXXVII, No. 8. The American Translators Association, Available at http://www.justrightcommunications.com/files/Moraes_Binoculars_p1_Aug08.pdf, accessed 29 January 2013.

Tavast, Arvi. 2008. The translator is human too: a case for instrumentalism in multilingual specialised communication. Dissertation. University of Tartu. Tartu: Tartu Ülikooli Kirjastus.

Vandepitte, Sonia. 2008. Remapping Translation Studies: Towards a Translation Studies Ontology. *Meta: Translator's Journal*, volume 53. Available at http://www.erudit.org/revue/meta/2008/v53/n3/019240ar.html, accessed 15 April 2013.

Vettik, Aime. 2011. Õiguskeele kaks arenguteed. *Õiguskeel* 2012/3. Available at http://www.just.ee/41256, accessed 12 January 2013.

Wallis, Julian. 2006. Interactive Translation vs Pre-translation in the Context of Translation Memory Systems: Investigating the effects of translation method on productivity, quality and translation satisfaction. MA thesis. University of Ottawa.

USED PROGRAMS

Programs discussed or used in this paper:

- AntConc 3.2.4w (Windows) 2011, available from: http://www.antlab.sci.waseda.ac.jp/software.html
- Festart Dictionary v2006.51 Professional, available from: Festart OÜ (Chapter 1)
- Microsoft Excel, available from: http://office.microsoft.com/en-gb/products/?CTT=97
- Microsoft Word, available from: http://office.microsoft.com/en-gb/products/?CTT=97
- SDL MultiTerm 2011, available from: www.sdl.com
- SDL MultiTerm Extract, available from: www.sdl.com
- SDL Trados Studio 2011, available from: www.sdl.com

APPENDICES

APPENDIX 1: LIST OF CAT TOOLS

The following list includes some most commonly used and known computer assisted translation tools relevant to the subject discussed in this paper. The list of categories is in random order, the lists of software or programs have been given in alphabetical order.

Word processors
AbiWord
Google Docs
KWord
LibreOffice Writer
LyX
Microsoft Word
Office Web Apps
OpenOffice Writer
WordPerfect

Spreadsheet applications
Corel Quattro Pro (WordPerfect Office)
Gnumetric
GNU Oleo
IBM Lotus Symphony
Kingsoft Spreadsheets
KSpread
Lotus 1-2-3 and other MS-DOS spreadheets
Microsoft Excel
Number by Apple Inc.
VisiCalc

Translation memory programs
Across
Déjà Vu
memoQ by Kilgray
OmegaT
SDL Trados Studio
Wordfast
WordServer

Terminology extraction and management software
LogiTerm
memoQ 5.0 by Kilgray
SDL MultiTerm, SDL MultiTerm Extract
SyncroTerm
Termex

Corpus-processing tools

AdTAT
AntConc
ApSIC Xbench
CorpusEye
GlossaNet/Unitex
KH Coder
Linguistic Toolbox
LogiTerm Pro
MonoConc
TextSTAT
WordSmith

APPENDIX 2.1: QUESTIONNAIRE (ENGLISH)

Translation of terms

Dear translator,
I am a student of Tartu University and am writing my thesis on means used to achieve consistent translation of terms. For the purpose of analysing the situation and receiving information important to translators, please answer the following questions.
You will remain anonymous.

1) Have you studied translation academically (count also unfinished studies)?

☐ Yes ☐ No

2) What is your translation experience (work or study)?

☐ Only for study purposes

☐ Mostly for study purposes

☐ Mostly for work

☐ Only for work

3) Assess the duration of your translation experience:

☐ Short - less than one year or some years in low volumes

☐ Medium - some years (1-5) in high volumes or many years (over 5) in low volumes

☐ Extensive - many years (over 5) in high volumes

4) Have you studied any of the computer programs you use for translating in an academic setting? (Select all that apply.)

☐ Yes, at university

☐ Yes, at school

☐ Yes, at a training course

☐ No

Please specify your answer in free form. E.g. How long, on what level, which programs were covered in the studies?

5) Do you feel you would need to or like to improve your skills and knowledge in using computer programs?

Translation case:
There is a 17-page translation order consisting of two MS Word documents. One is a contract, the other, general contract conditions annexed to the contract. You will have three full days before you

are due to submit the finished translation. That will leave you a little over five pages per day to translate.
Please answer the following questions pretending you have taken the order, but keeping in mind how you have gone about with similar work previously.

6) Will you translate the documents:

- ☐ Using translation memory software (e.g. SDL Trados, WordFast)
- ☐ Using translation memory software with terminology software (e.g. SDL Trados Studio with SDL MultiTerm)
- ☐ In MS Word opening a new file for the translation
- ☐ In MS Word writing on top of the source language text replacing it

Should you not use any of the named, please specify:

7) Please explain in short, why will you translate the documents using the software/program you have selected in question 6.

8) Will you use any of the following to aid translating? (Select all that apply)

- [ ] SDL MultiTerm or similar termbase
- [ ] AntConc or similar concordance program
- [ ] Self-made termbase for that particular translation (in MS Word, MS Excel or other electronic means)
- [ ] Self-made termbase for that particular translation (on paper)
- [ ] Replace-function

9) Before starting to translate, will you read the text?

- [ ] Yes, I will read it through carefully
- [ ] Yes, I will skim the text through
- [ ] No, I will only have a quick look and start work

If yes, do you mark recurring terms in the text in any way (underlining, highlighting, writing them out, etc.)?

- [ ] Yes
- [ ] No

If no, do you use any software (e.g. SDL MultiTerm Extract; AntConc) to find terms in the text?

- [ ] Yes
- [ ] No

Please specify!

10) When coming across a term in text while translating and you need to take 5 or more minutes to find an equivalent for it, what do you do with the term?

- [ ] Translate it in the one place and memorize the answer
- [ ] Translate it in the one place and write it down for yourself (e.g. in MS Word, MS Excel or on paper)
- [ ] Translate it and apply the translation throughout the text (e.g. using the Replace-function)
- [ ] Translate it and add it to a termbase (e.g. SDL MultiTerm)

If you feel the provided answers have restriced you too much and you need to express your opinion, please do so here:

Any other kind of feedback is welcome on: piret99j@ut.ee

APPENDIX 2.2: QUESTIONNAIRE (ESTONIAN)

Terminite tõlkimine

Lp tõlkija

Olen Tartu Ülikooli üliõpilane ning kirjutan magistritööd ühtsuse tagamisest terminite tõlkimisel. Selleks, et saaksin olukorda uurida ning tõlkijatega olulist teavet jagada, palun teil vastata käesolevale küsimustikule.

Teie anonüümsus on tagatud.

1) Kas olete õppinud tõlkimist akadeemilises keskkonnas (arvestada ka pooleliolevaid õpinguid)?

☐ Jah ☐ Ei

2) Milline on teie tõlkekogemus (õpingud või töö)?

☐ Ainult õpingutega seotud

☐ Peamiselt õpingutega seotud

☐ Peamiselt töö

☐ Ainult töö

3) Hinnake oma tõlkekogemuse kestvust:

☐ Lühike - alla ühe aasta või mõned aastad väikses mahus

☐ Keskmine - mõned aastad (1-5) suures mahus või mitu aastat (üle 5) väikses mahus

☐ Pikk - mitu aastat (üle 5) suures mahus

4) Kas olete õppinud oma igapäevatöös kasutatavaid arvutiprogramme akadeemilises keskkonnas? (Valige kõik sobivad!)

☐ Jah, ülikoolis

☐ Jah, koolis

☐ Jah, koolitusel või kursusel

☐ Ei

Palun täpsustage vastust vabas vormis. Nt: mitu aastat, millisel tasemel, milliseid programme kasutama õppisite.

5) Kas sooviksite oma arvutiprogrammide kasutamise oskust tõsta? Kas see oleks teie igapäevatöö tõhustamiseks vajalik?

Tõlkejuhtum:
Olete vastu võtnud 17-leheküljelise tõlketellimuse. Töö koosneb kahest MS Word dokumendist. Üks on leping, teine lepingule lisatud lepingu üldtingimused. Teil on kolm täispäeva tõlke tegemiseks – seega natuke üle viie lehekülje päevas.
Palun vastake järgnevatele küsimustele, pidades silmas seda tõlkejuhtumit ning arvestades, kuidas tavaliselt kirjeldusele vastavaid töid teinud olete!

6) Kas tõlgite dokumendid:

- [ ] Kasutades tõlkemäluprogrammi (nt SDL Trados, WordFast)
- [ ] Kasutades tõlkemäluprogrammi koos terminitöötlustarkvaraga (nt SDL Trados Studio ja SDL MultiTerm)
- [ ] Programmis MS Word kirjutades tõlge uude faili
- [ ] Programmis MS Word kirjutades tõlge originaalteksti asemele

Kui te ei kasuta tõlkimiseks ühtegi nimetatud variantidest, palun täpsustage:

7) Palun selgitage lühidalt, miks tõlgite dokumendid kasutades küsimuses 6 valitud tarkvara/ programmi!

8) Kas kasutate tõlkimisel abiks järgnevaid programme või võimalusi? (Valige kõik sobivad!)

- [ ] SDL MultiTerm või sarnane terminibaas
- [ ] AntConc või sarnane konkordantsiprogramm
- [ ] Ise konkreetse tõlke jaoks koostatud terminibaas (programmis MS Word, MS Excel või muul elektroonilisel viisil)
- [ ] Ise konkreetse tõlke jaoks koostatud terminibaas (paberil)
- [ ] Asendusfunktsioon `Replace`

9) Kas loete teksti enne tõlkima asumist?

- [ ] Jah, loen põhjalikult läbi
- [ ] Jah, tutvun teksti sisuga, aga mitte väga põhjalikult

☐ Ei, viskan vaid pilgu peale ja hakkan kohe tööle

Kui loete, kas märgite tekstis korduvaid termineid mingil viisil (alla joonides, värvides, kirjutades need välja vms)?

☐ Jah ☐ Ei

Kui ei loe, kas kasutate tekstist terminite leidmiseks mingit tarkvara (nt SDL MultiTerm Extract; nt AntConc)?

☐ Jah ☐ Ei

Palun täpsustage!

10) Kui kohtate tõlkimise ajal tekstis terminit, mille vaste otsimiseks ja selle üle otsustamiseks kulutate 5 minutit või rohkem, mis te sellega teete?

☐ Kirjutan selle tõlke hetkel käsil olevasse lausesse ning jätan vaste meelde

☐ Kirjutan selle tõlke hetkel käsil olevasse lausesse ning märgin selle vaste üles (nt programmi MS Word, MS Excel või paberile)

☐ Kirjutan selle tõlke hetkel käsil olevasse lausesse ning ka kõikidesse teistesse kohtadesse, kus termin tekstis esineb (nt kasutades asendusfunktsiooni "Replace")

☐ Kirjutan selle tõlke hetkel käsil olevasse lausesse ning lisan selle terminibaasi (nt SDL MultiTerm)

Kui teile tundub, et pakutud variandid on teid liigselt piiranud või ei lase teil kuvada täispilti oma tööprotsessist või valikutest, palun kasutage allolevat lahtrit kommentaarideks.

Igasugune muu tagasiside on teretulnud aadressil: piret99j@ut.ee

APPENDIX 3: SDL TRADOS STUDIO 2011 SOFTWARE DEMONSTRATION SCREEN SHOTS

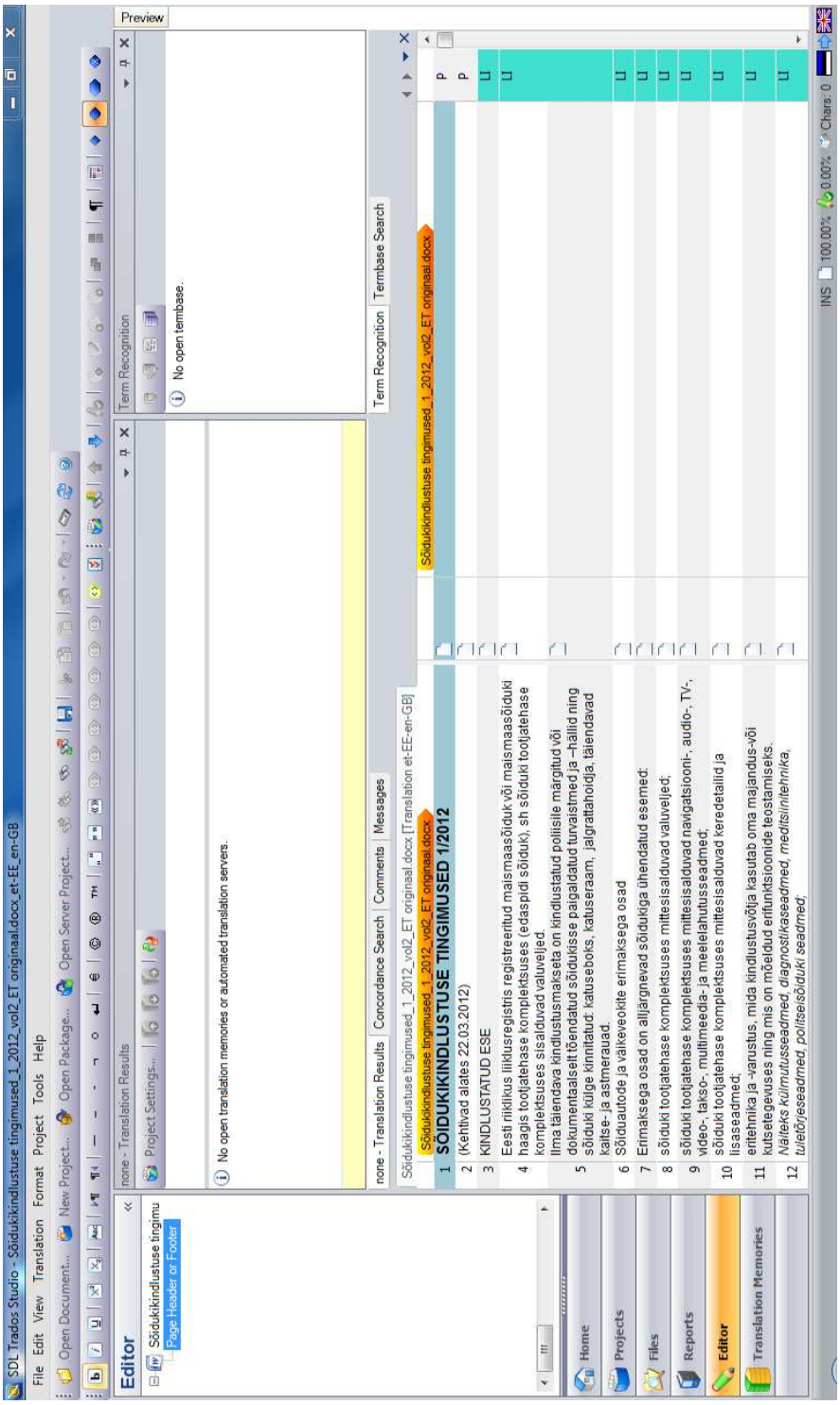


Figure 16: File to be translated opened in SDL Trados Studio

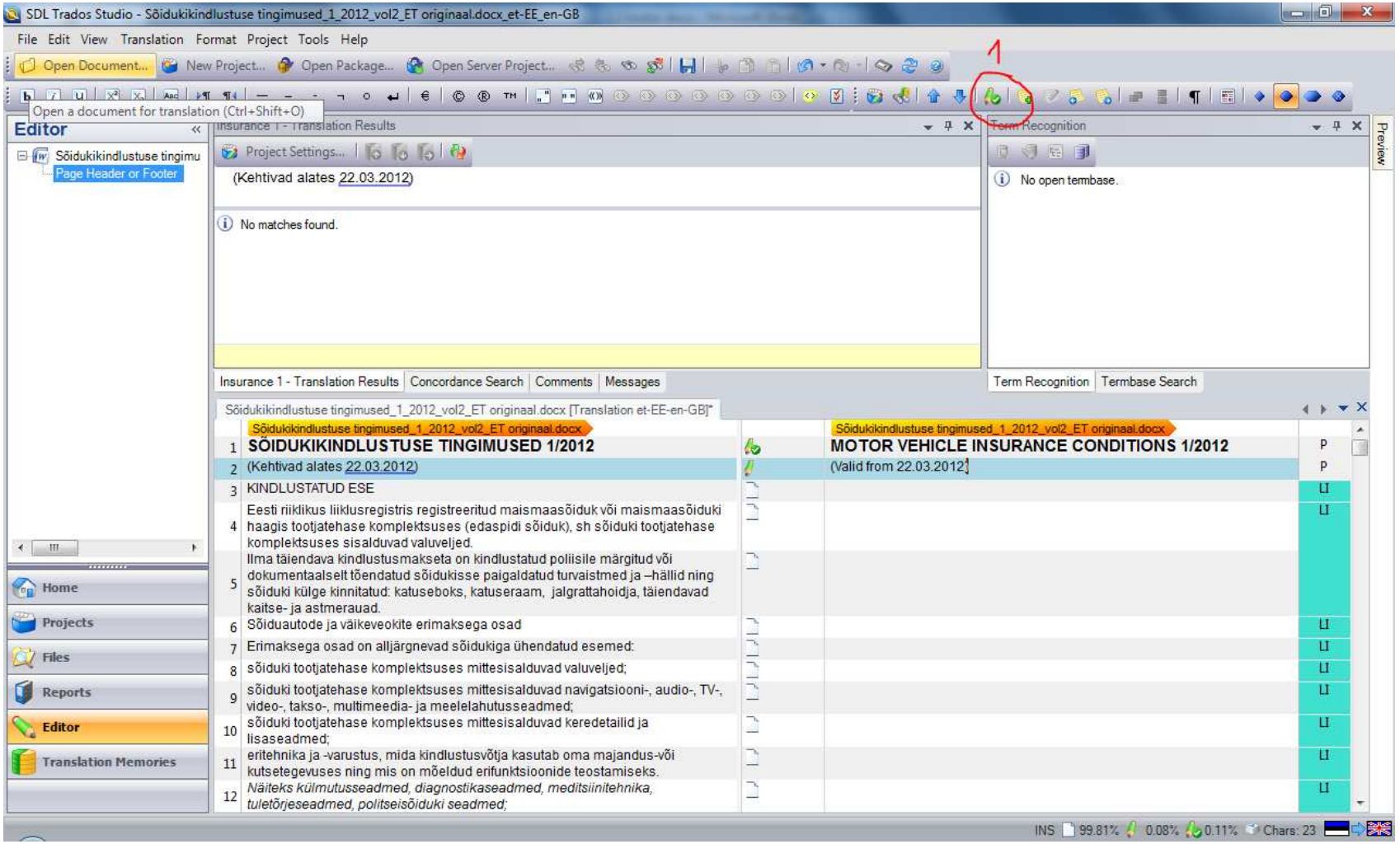


Figure 17: Confirming the translated section in SDL Trados Studio

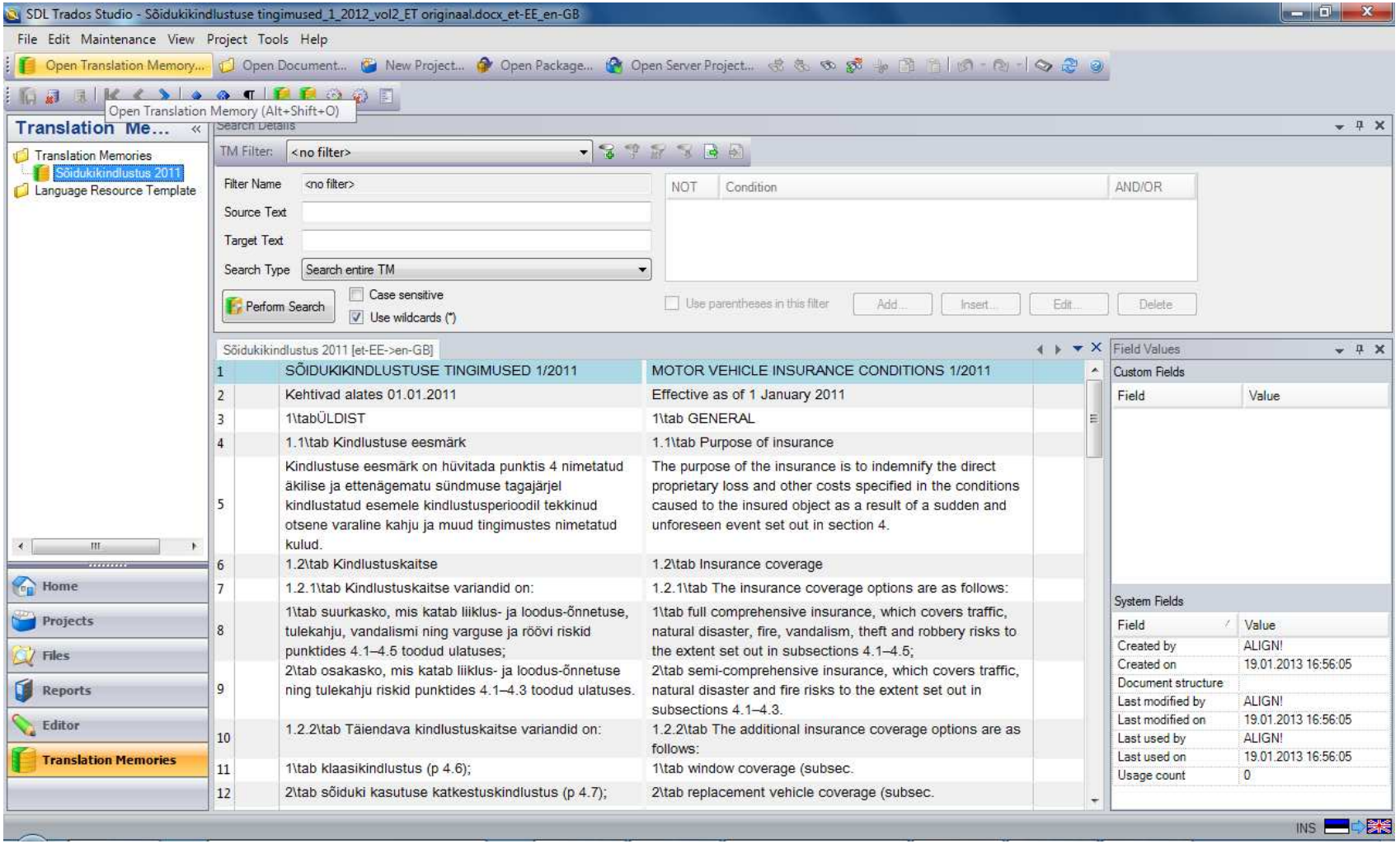


Figure 19: Aligned translation memory in SDL Trados Studio

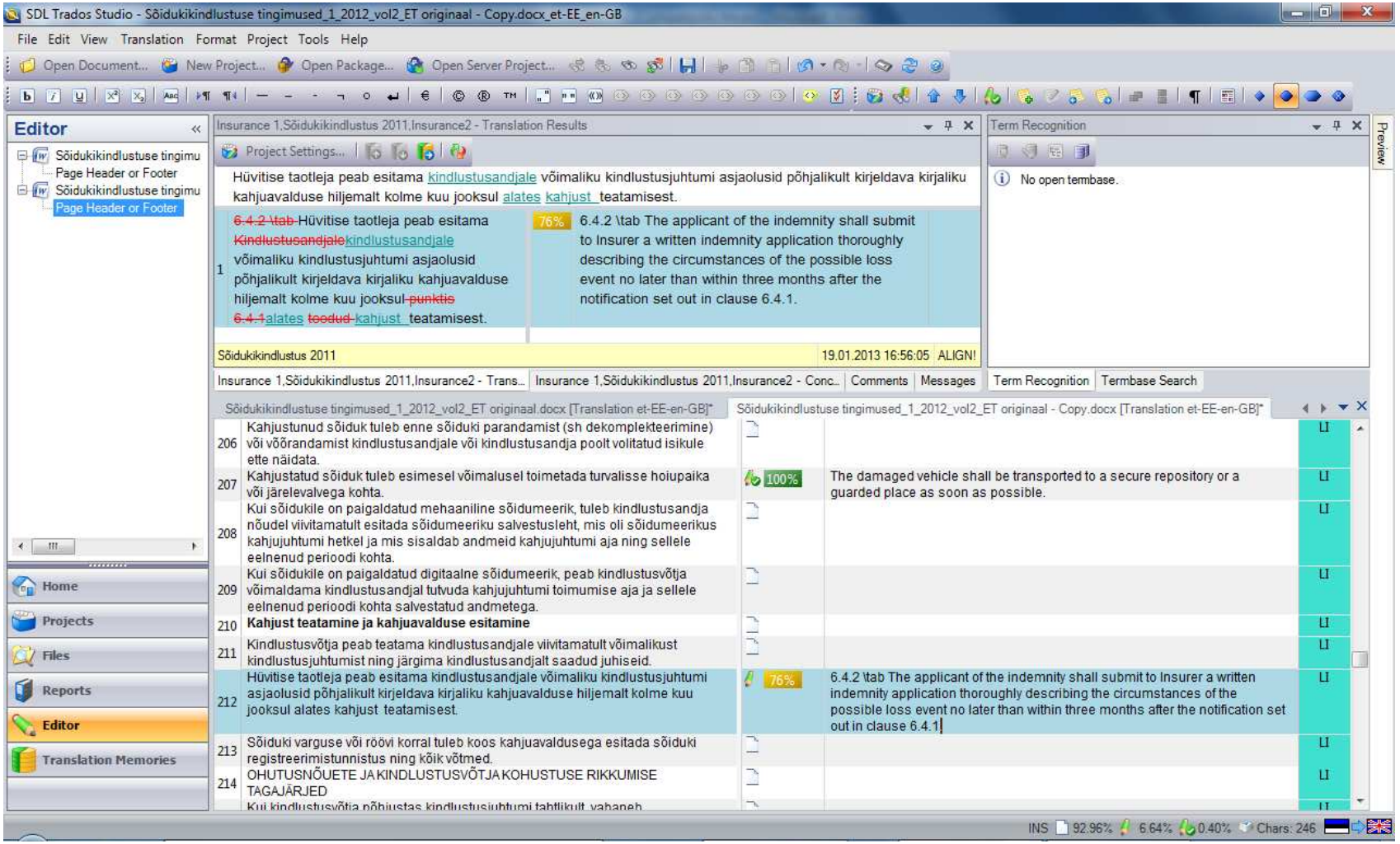


Figure 20: Translation memory results in SDL Trados Studio

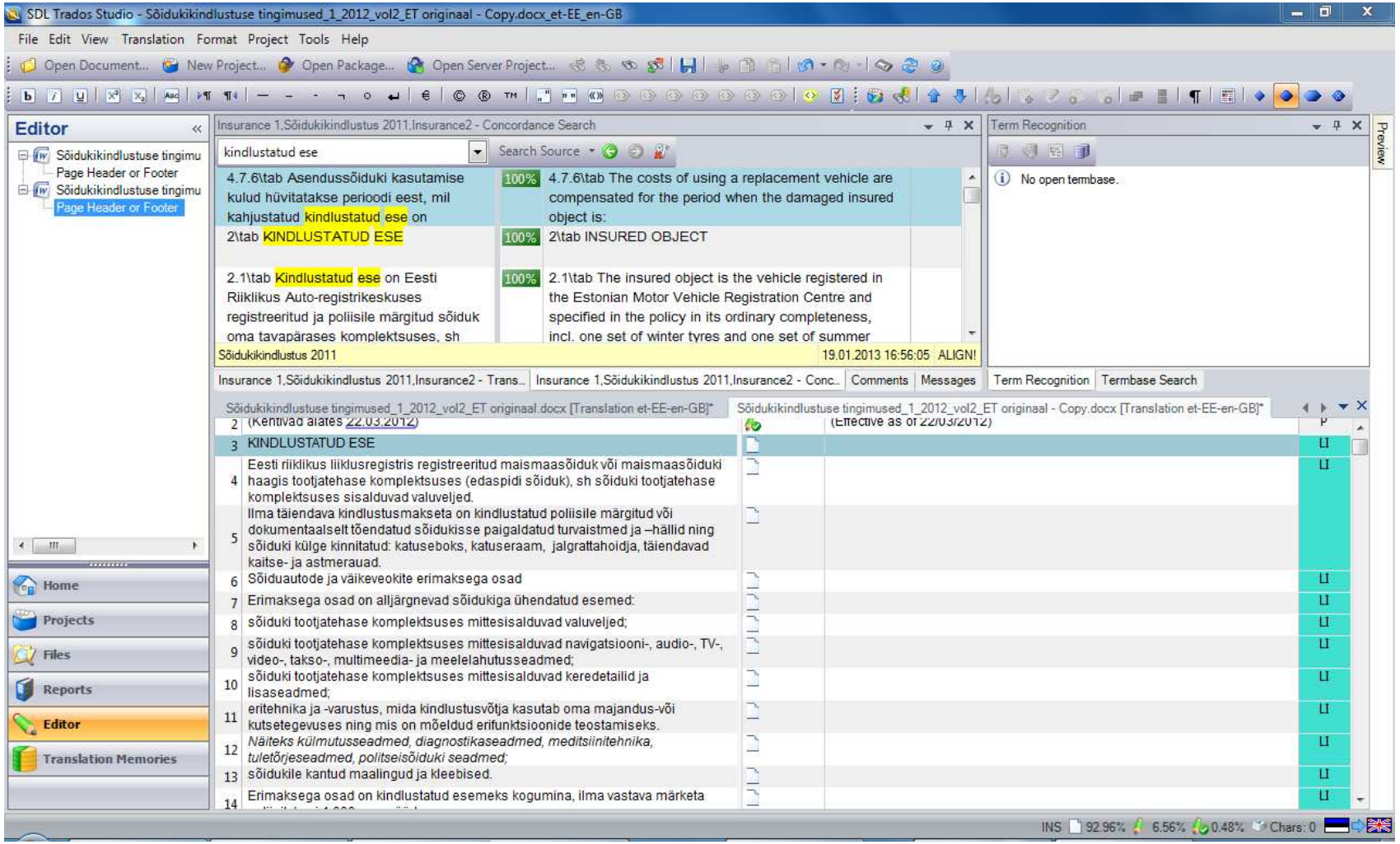


Figure 21: Concordance search in translation memory files in SDL Trados Studio

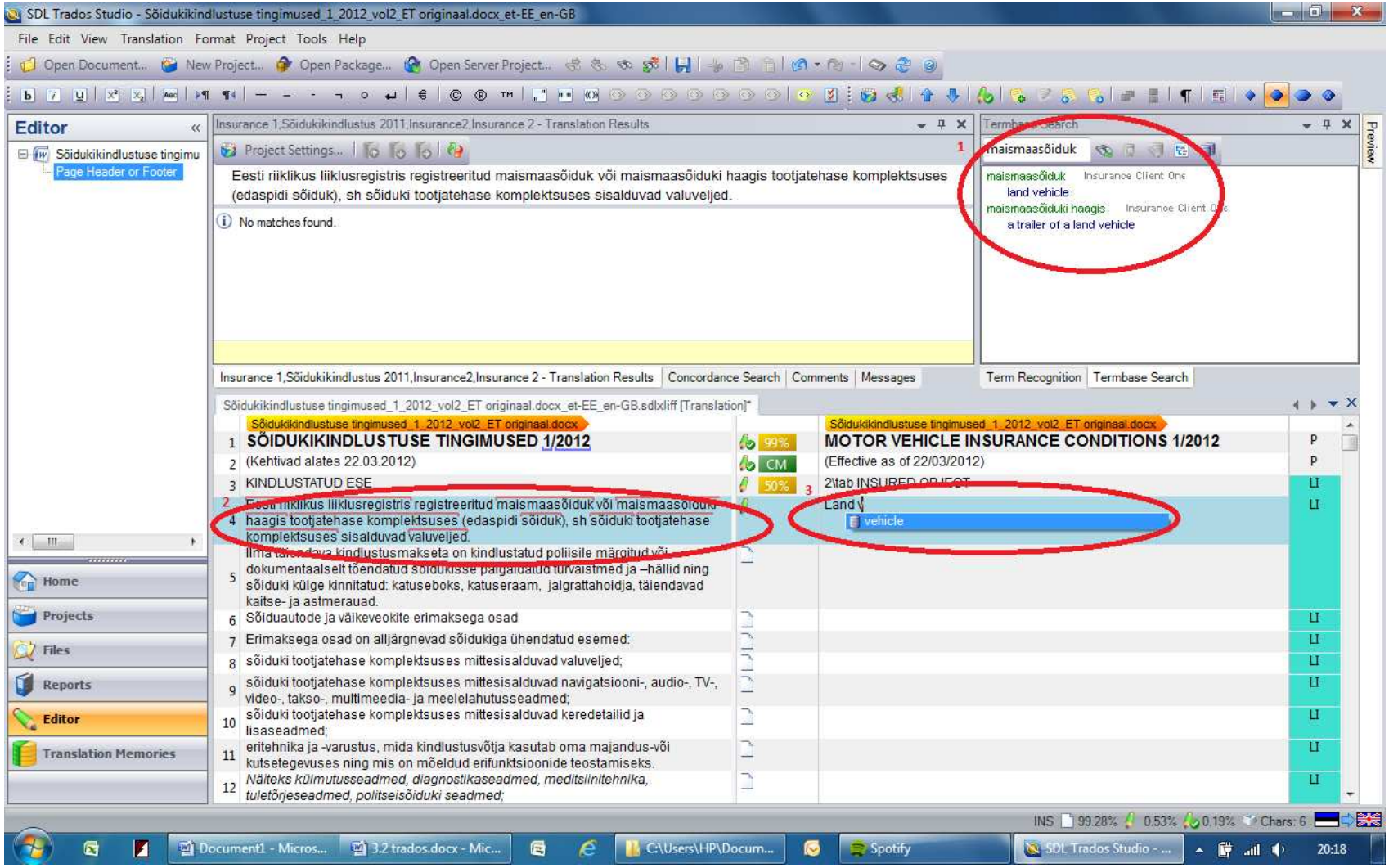


Figure 24: Using a termbase in SDL Trados Studio

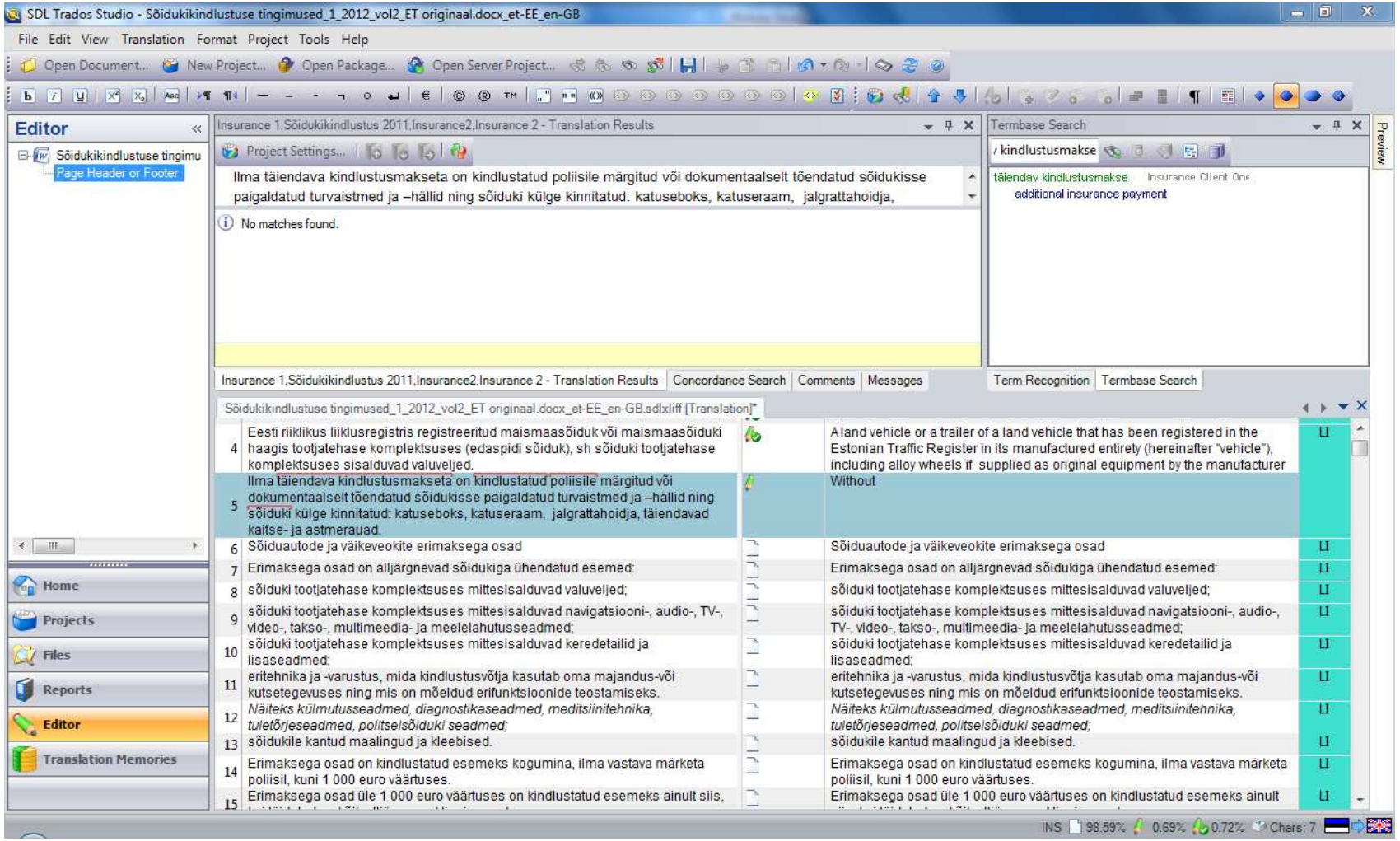


Figure 25: ST copied into TT fields in SDL Trados Studio

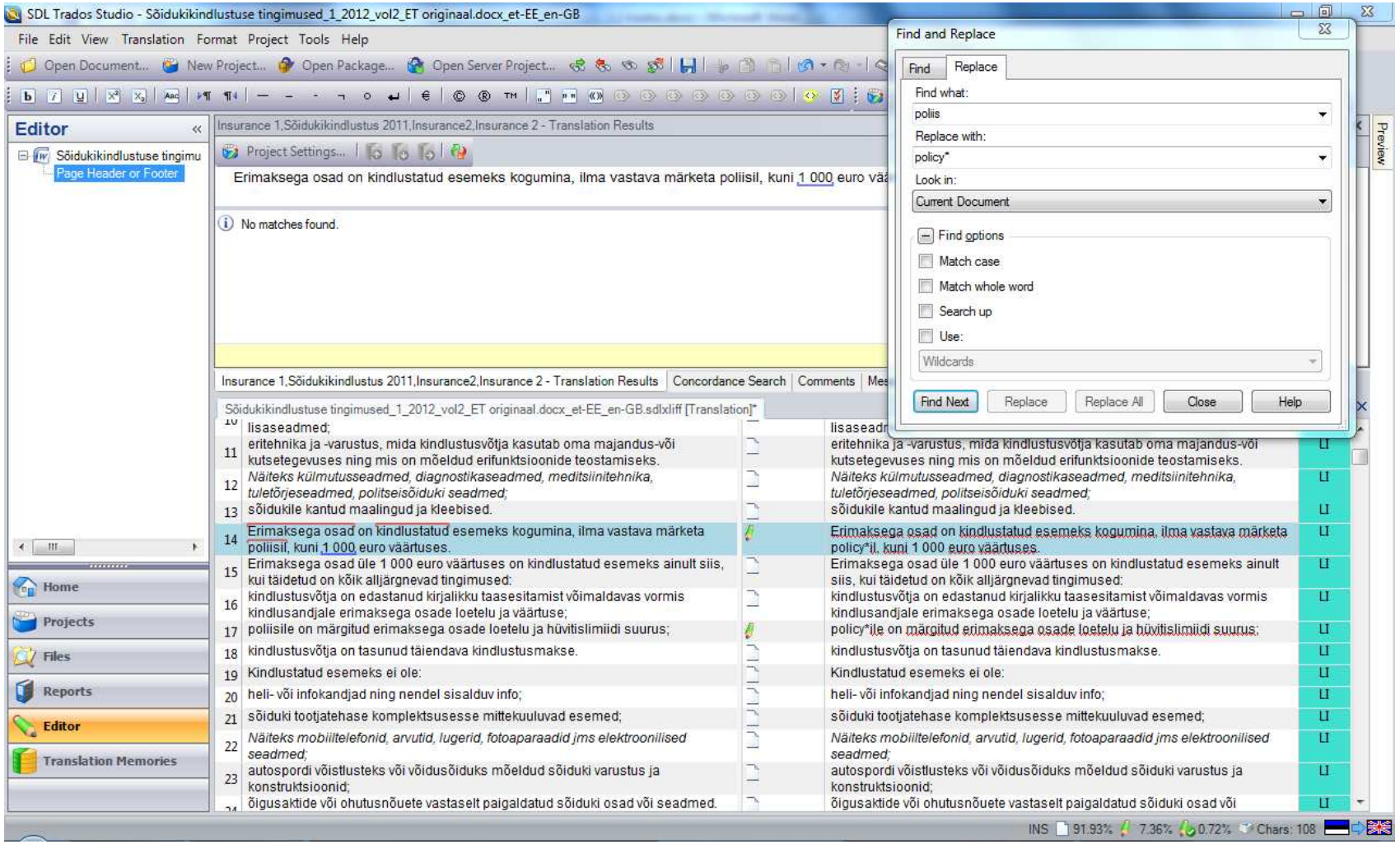


*Figure 26: Applying terms using **Replace** function in SDL Trados Studio*

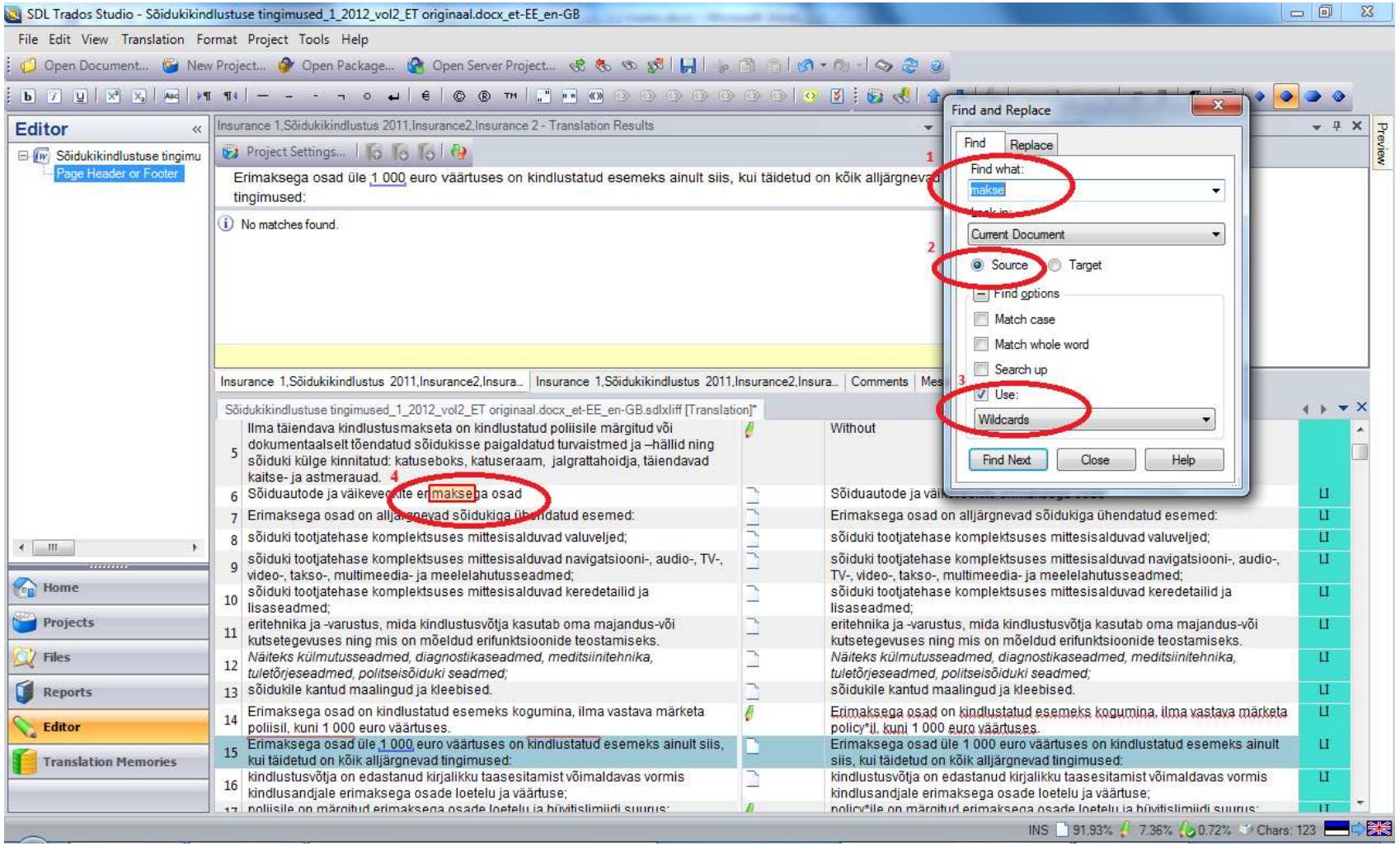


Figure 27: Finding terms in the ST in SDL Trados Studio

RESÜMEE

TARTU ÜLIKOOL
GERMAANI, ROMAANI JA SLAAVI FILOLOOGIA INSTITUUT

Piret Jõgis

Employing technical means for ensuring consistent translation of recurring terms
Tehniliste vahendite rakendamine korduvate terminite ühtselt tõlkimiseks

Magistritöö
2013
Lehekülgede arv: 110 (koos lisadega)

Annotatsioon:

Tõlked, kus tekstis esinev termin on sihtkeeles saanud mitu vastet, on nii klientidele kui tõlke teistele kasutajatele häirivad, vahel ka segadust tekitavad. Selleks, et selliseid vigu ei tekiks, on vaja, et tõlkijad oskaks probleemile piisavalt tähelepanu pöörata ja teha oma tööd selliselt, et sarnaseid vigu ei tekiks. Tänapäeval, mil tõlkijad kasutavad oma töös tehnilisi abivahendeid, on asjakohane uurida, kas ja kuidas saavad kasutatavad programmid tõlkijat vastavalt abistada.

Seetõttu püstitati magistritöö eesmärgiks leida viisid, kuidas erinevad tarkvaraprogrammid saavad aidata professionaalsetel tõlkijatel tagada terminite läbivalt ühtselt tõlkimist.

Töö koosneb kolmest osast, millest esimene toob välja korduvate terminite käsitlemise võimalikke meetodeid teoreetilisest aspektist. Samuti tutvustatakse selles osas erinevaid tõlkimisel kasutatavaid programme. Selleks, et saada rohkem teavet tõlkijate töömeetodite kohta, viidi töö osana 2013. aasta veebruaris läbi küsitlus, millele paluti vastata tegutsevatel tõlkijatel. Töö teises osas kirjeldatakse küsitlust ja antakse ülevaade selle tulemustest. Kolmandas osas on küsitlustulemustele toetudes valitud levinumad tõlkeabiprogrammid, mida vaadeldakse korduvate terminite ühtse tõlkimise vaatenurgast, üritades leida viise, kuidas saavutada võimalikult vähese vaevaga parim tulemus. Lisaks tõlkijate poolt juba kasutatavatele programmidele käsitleti ka võimalust lisada korduvate terminite tõlkimisühtsuse tagamiseks tõlkija tööriistade hulka konkordantsiprogramm.

Töö tulemusena leiti mitmeid viise, kuidas tõlkija saab korduvate terminite tõlkimisel programme rakendada. Küsitluse tulemusena saadi teada, milliseid programme tõlkijad oma töös kasutada eelistavad ja millisel määral neid terminitöös kasutatakse. Ka leiti, et terminite tõlkimisel usaldatakse suurel määral oma mälu. Töös leiti võimalusi, kuidas need mäluülesanded tarkvaraprogrammidele üle anda selliselt, et tõlkija peaks oma harjunud töömeetodeid võimalikult vähe muutma. Vastavad soovitused sisalduvad tarkvara demonstratsioonides. Töö eesmärk anda tõlkijatele ideid, kuidas tõhustada oma tööd korduvate terminite ühtsel tõlkimisel, sai täidetud.

Märksõnad: kirjalik tõlge, tõlkimine, tõlkeabiprogrammid, tõlkimisel kasutatav tarkvara, tõlkemälu, tõlkemäluprogrammid, SDL Trados Studio, SDL MultiTerm, AntConc, terminite ühtne tõlkimine, terminibaas.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, _____Piret Jõgis,_____

(*autori nimi*)

(sünnikuupäev: 17.02.1983)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
_____”Employing technical means for ensuring consistent translation of recurring terms”
(*lõputöö pealkiri*)

mille juhendaja on _____Piret Rääbus_____,

(*juhendaja nimi*)

1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace´i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, „20” mail 2013